THE

AMERICAN

JOURNAL OF PSYCHOLOGY

EDITED BY

G. STANLEY HALL,

E. C. SANFORD,
Clark University.

AND

E. B. TITCHENER,
Cornell University.

WITH THE CO-OPERATION OF

F. ANGELL, Stanford University; H. BEAUNIS, Universities of Nancy and Paris; I. M. BENTLEY, Cornell University; A. F. CHAMBERLAIN, Clark University; C. F. HODGE, Clark University; A. KIRSCHMANN, University of Toronto; O. KUELPE, University of Würzburg; W. B. PILLSBURY, University of Michigan; A. D. WALLER, University of London; M. F. WASHBURN, Vassar College.

VOL. XV.

CLARK UNIVERSITY, WORCESTER, MASS.
LOUIS N. WILSON, Publisher.

1904.

R.H.

BF 1
A51
v. 15

COPYRIGHT, 1904, BY G. STANLEY HALL.

OHIO STATE
UNIVERSITY

COMMONWEALTH PRESS,
OLIVER B. WOOD, PROPRIETOR,
WORCESTER, MASS.

TABLE OF CONTENTS.

BENJAMIN RICHARD ANDREWS,
Auditory Tests, ... 14-56

L. D. ARNETT,
The Soul—A Study of Past and Present Beliefs, 121-200; 347-382

I. MADISON BENTLEY,
Professor Cattell's Statistics of American Psychologists, 102-103

I. M. BENTLEY AND E. B. TITCHENER,
Ebbinghaus' Explanation of Beats, . . . 62-71

GIOVANNI CHIABRA,
The Tendencies of Experimental Psychology in Italy, . 515-525

C. E. GALLOWAY,
The Effect of Stimuli upon the Length of Traube-Hering Waves, ... 499-512

ROBERT H. GAULT,
A Sketch of the History of Reflex Action in the Latter Half of the Nineteenth Century, ... 526-568

BERTHA KILLEN,
The Effects of Closing the Eyes upon the Fluctuations of the Attention, ... 512-514

F. KUHLMANN,
Experimental Studies in Mental Deficiency: Three Cases of Imbecility (Mongolian) and Six Cases of Feeble-Mindedness, ... 391-446

ROBERT MACDOUGALL,
Facial Vision: A Supplementary Report, with Criticisms, 383-390

W. P. MONTAGUE,
A Theory of Time-Perception, ... 1-13

W. B. PILLSBURY,
Studies from the Psychological Laboratory of the University of Michigan, ... 499-514

JAMES P. PORTER,
A Preliminary Study of the Psychology of the English Sparrow, ... 313-346

THEODATE L. SMITH,
The Psychology of Day Dreams, ... 465-488

60700

C. SPEARMAN,
The Proof and Measurement of Association between Two Things, ... 72-101
"General Intelligence," Objectively Determined and Measured, ... 201-293

GUSTAV SPILLER,
The Problem of the Emotions, ... 569-580

H. C. STEVENS,
A Simple Complication Pendulum for Qualitative Work, 581

E. B. TITCHENER,
Some New Apparatus, ... 57-61

E. B. TITCHENER and I. M. BENTLEY,
Ebbinghaus' Explanation of Beats, ... 62-71

GUY MONTROSE WHIPPLE,
Reaction-Times as a Test of Mental Ability, . . 489-498

ANONYMOUS,
Nocturnal Emissions, ... 104-107

ERRATA, ... 120

LITERATURE, ... 108-120; 294-312; 447-462; 582-596

NOTES, ... 312; 463-464

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded by G. STANLEY HALL in 1887.

VOL. XV. JANUARY, 1904. No. 1.

A THEORY OF TIME-PERCEPTION.

By W. P. MONTAGUE, Columbia University.

I.

THE SPECIOUS PRESENT.

The first question to be faced in a study of time-perception is the question of the "specious present;" for without the consciousness of an extended segment, or period, of past time it would be impossible to perceive either duration or succession. This problem of the specious present is not only the most important,—it is also the most perplexing of the many problems of our time consciousness. It is a case in which sense-perception presents as an actuality what Reason must regard as an impossibility. The present of metaphysical or conceptual time is a point, separating past and future; the present of psychological time—the specious present—is a continuous segment extending appreciably into the past. We cannot hope to solve this antinomy by violating Reason; we must not accept the presence of what is no longer present as a reality. Sense must be subordinated to Reason, and the specious present must be regarded as specious, as an illusion which is somehow explicable on the assumption that the real present is a point. The problem may then be stated as follows:

How is it that at any one moment there can appear to be present several moments?

II.

THE MEASUREMENT OF TIME.

Time is the form of change, and the amount of change is the measure of the time which is perceived to have elapsed during the change, precisely as the size of a body is the measure of the

space actually or possibly filled by the body.[1] Taken in abstraction from all content, space and time, if not nothing, are at least quite without determinate form and size. As space and time are measured by what fills them, so, too, is the concrete filling measured by comparison with other concrete filling. A body is large or small according to the smallness or largeness of a second body selected as standard of measure. A change is great or small according to its relation to some other change which we take as a basis of comparison. Magnitude, whether conceptual or perceptual, is intrinsically relative.

III.

The Subjective and Objective Aspects of a Psychosis.

Every psychosis has two distinguishable but inseparable aspects, the subjective and the objective. The subjective element or "knowing thought" is the whole system of conscious contents taken collectively and including the incoming content, while the latter is the *object* of the (normally prospective) act of attention. A content is perceived as an object only in so far as it is projected upon a subjective background of pre-existing states. In Herbartian terms, we may describe every psychosis as the assimilation of an entering sensation-mass by a receiving apperception-mass.

I believe that the solution of the problem of the specious present depends on our realizing that the subjective and objective elements of a psychosis possess differing rates of change.

What is sometimes called the Law of Relativity is an expression of the fact that in consciousness every content is determined not only by its stimulus, but also by every other content. A change in one state is necessarily attended by a change in the other states, both individually and collectively. If we are right in regarding the knowing thought or subjective aspect

[1] It might be objected that the experience of unequal amounts of change occurring in equal times is in flat contradiction to the statement that time is measured by its contents. When, however, we seek for the grounds on which the two unequally filled times are judged equal, we see that their estimated equality arises from our comparing them, not with an absolute time, but with other changes of content, either those of our subjective consciousness or those of a third object, such as a clock. That absolute or empty time does not enter as a factor into the perceptual estimate of time is further shown by the fact that when equal time-intervals containing unequal quantities of change are perceived separately and without relation to a third object, they are, as a matter of fact, felt to be unequal. And the estimated inequality of their times is in proportion to the perceived inequality of their contents. Notwithstanding the uselessness of absolute time as a standard of perceptual measurement, the idea has a certain validity. The absolute clock is the totality of things, viewed as a single system.

of a psychosis as the totality of existing states, considered as forming a system having a single structure, or configuration, it will be clear that the changes induced in the form of that totality by a change in one of its parts will usually be less than the change in that part. In a word, a change in one conscious state produces a change (though a less change) in the system of states; and other things equal, the larger the apperception-mass, the smaller the changes produced in it by the changes in any one of its sensory aspects. The perception of a change is a changing perception, but a perception that changes less rapidly than the change that is perceived.

It cannot be too strongly emphasized in this connection that the increment of change produced in a conscious state,—or apperceptive system of states,—while it varies directly with the stimulus causing it, also varies inversely with the content in which it is caused. Thus if s be a conscious state and Δs be its increment of change, Δs will be large in proportion as s is small. In this respect, and, so far as I can see in this respect only, do increments of *change* differ from the increments of *substance*, ordinarily considered in the calculus.

There is a second point to which I would especially call the reader's attention before proceeding further. I have spoken, and shall speak throughout, of conscious states as quantities, expressible by quantitative symbols, and subject to quantitative laws: yet, nevertheless, these quantitative expressions can be, *and perhaps should be*, interpreted as applying to the *physiological concomitants* of conscious states. The reader who is sensitive on this subject can substitute the physical for the psychical interpretation of the quantitative symbols, without changing the meaning of the argument at any point.

Let Δo symbolize the amount of change or alteration in the objective content o produced in any period of time Δt, and let Δs symbolize the resulting change produced in the subjective aspect of consciousness during the same time. Then $\frac{\Delta o}{\Delta s}$ will represent a change in the objective as compared with the change in the subjective element in the time Δt. As Δt is made to decrease without limit, Δo and Δs will correspondingly decrease, but the fraction $\frac{\Delta o}{\Delta s}$ will not necessarily decrease, but will either approach, or if the rate of change be uniform, will maintain the finite value $\frac{do}{ds}$. Now this derivative of the objective change with regard to the change of the subjective element is a finite quantity, but one that is realized at each infinitesimal moment of time. If we were right in affirming in the preceding section that perceptual space and time magnitudes are essentially

relative matters, determined not by the "absolute" size of the magnitude to be measured but by the ratio of that magnitude to the unit of measure, it will be evident that the amount of objective time or change which appears to be present at any one moment will be measured by its ratio to the subjective change which accompanies it. And if the subjective unit of measure be infinitesimal, the objective change-magnitude may be also infinitesimal and yet may *appear* finite. In short the infinitesimal time-change which occurs at the unextended metaphysical present will be appreciated as an extended psychological present, provided it be measured in terms of an infinitesimal of different order.

That $\triangle s$, the change in the subjective totality of conscious contents, is less or "slower" than $\triangle o$, the change in that phase of the totality selected as object by the act of attention, has been shown: consequently the symbol $\frac{do}{ds} = \lim_{\triangle t = o} \left[\frac{\triangle o}{\triangle s}\right] > 1$ may be accepted as properly representative of the "specious" or "extended" present; *and the latter is from this point of view seen to be compatible with the real or metaphysical present of conceptual time.*

IV.

Duration and Succession.

At the outset of our discussion it was stated that the perception of the specious present was prerequisite to the perception of duration and succession. Now introspection assures us that the specious present never maintains the same value from moment to moment. It is always either expanding or contracting. *But an increase in the specious present means an increase in the period of past time which appears at present, that is to say an increase in the time during which the events filling it have been perceived.* The objects or events, therefore, which occupy the expanding specious present are felt as "enduring," while conversely those that fill the contracting specious present are felt as "passing" or "succeeding." This alternate increase and decrease of $\frac{do}{ds}$ implies of course a corresponding alternation in the sign of the second derivative. When an object o is perceived as enduring, $\frac{d^2o}{ds^2}$ is positive; at the moment when o is succeeded and displaced in consciousness by some other object, $\frac{d^2o}{ds^2}$ is negative.

V.

"Time Dragging" and "Time Flying."

The second derivative of the objective with regard to the

subjective change not only indicates by its *sign* the enduring or passing phase of its contents, but by its *magnitude* it marks the *seeming rate of the time-flow itself*. A small or large change-rate (whether increasive or decreasive) in $\frac{do}{ds}$ means a correspondingly small or large value of $\frac{d^2o}{ds^2}$. In view of the fact previously mentioned that $\frac{s}{o}$ varies as $\frac{do}{ds}$, it is interesting to note that, as might be expected, the first and second derivatives also vary inversely. The longer an event endures, *i. e.*, the greater the value of $\frac{do}{ds}$, the less the rate of its change, and the less the value of $\frac{d^2o}{ds^2}$. The further past an event seems, the more slowly does it appear to increase in pastness, while conversely, the events of the present are those which seem to fly by most rapidly.

This would seem to afford a clue to the curious alterations which we experience in the *rate* of time-flow. For it is obvious that when new objects continuously distract the attention, the specious present is constantly being destroyed and made anew; and, as we saw, this enforced smallness of $\frac{do}{ds}$ would mean a correspondingly large value of $\frac{d^2o}{ds^2}$. The seeming rate of the time-flow is great in proportion to the number of new distractions. And conversely when the mind is without internal or external distractions, when there is nothing to think about and nothing to do, the specious present with its single content of *ennui*, *i. e.*, the consciousness merely of being conscious, grows to an enormous size with a consequent decrease in $\frac{d^2o}{ds^2}$, a decrease, that is to say, in the apparent rate of the time-flow. Time here is felt to *drag* as in the presence of distraction it is felt to *fly*.

VI.

The Rhythmic Character of Consciousness.

We have only to reflect upon the nature of the specious present as already described to understand the intrinsically rhythmic character of consciousness.

Let us suppose a mind to experience a series of sensations o_1, o_2, o_3, o_4, in moments t_1, t_2, t_3, t_4. The specious present at t_2 has a duration of two moments and contains the image-trace of o_1, and also the sensation o_2; and the specious present at t_3 has a duration of three moments and contains in addition to the sensation o_3, the images of o_1 and o_2. From the persistence of images the specious present tends to increase not only in "length" but in contents. This increase is not in a finite consciousness unlimited, but depends upon the degree to

which the accumulating contents are harmonized in a single meaning.

In consciousness as we know it, each content is to some degree a *rival* of its fellows. Of course it is also true that different contents to some extent support and enhance one another, and the law (first formulated, I believe, by Fechner) that the intensity of attention varies inversely as its extent is only a half truth. That is to say, it is only true for psychoses in which the "competition" of contents tends to become stronger than their "co-operation." The law is in general, however, rather more true than false, because as the number of elements in a system increases, co-operation becomes less easy and antagonism less difficult. And as the sweetest toned instrument will, if too loudly or too variously excited, reach a point at which the dissonance of the over-tones will exceed their harmony, so with the specious present there inevitably comes a time when the manifoldness of its contents makes their unity no longer capable of transcending their variety.[1]

Returning to our example we will suppose that in the mind which has experienced o_1, o_2, and o_3, the specious present has reached the critical point at which the further dispersal of attention demanded by a new content will result in its dissolution. The new content o_4 enters; the specious present ceases its slow expansion and begins rapidly to contract. The images of o_1, o_2 and o_3 crumble away and the attention thus liberated is caught by the strongest bidder which, in our case, is the present sensation o_4. On this as foundation a new specious present is rapidly built up, containing as a halo or "fringe" the shade of its parent. It is important to bear in mind the fact that the manner in which our specious presents are destroyed is different from the manner in which they are rebuilt. The old contents do not flow out as the new contents flow in. The events which are born singly die in cycles.

A rhythmic series of experiences affects us strongly and holds our attention because its objectively recurring periods harmonize with the subjectively recurring contractions and expansions of the specious present, each series being thus reinforced by the other. Subjectively viewed, the rhythmic character of consciousness is due, as we have seen, to the fact that each specious present contains in addition to the birth-seeds of new presents the seed of its own decay. Like other live things our time-

[1] Professor Loeb (*Comparative Physiology of the Brain*, Ch. XIX) holds that the intensity of a mental state is diminished by the simultaneous presence of other states, when and only when the underlying physiological processes are *aperiodic*. Simultaneous periodic processes so far from diminishing may actually enhance each others' intensity.

consciousness is rhythmic, because like them it is born with the canker of death at its heart.

VII.

MEMORY.

So far we have been dealing with the immediate apprehension of the just past. We have now to discuss the bearing of the theory upon the consciousness of the more remote past. Memory proper, or recall, is differentiated from the consciousness of the just past, by the fact that in the latter the event has been continuously present to consciousness since the time of its occurrence, while in recall, as the word indicates, the event remembered has been absent from consciousness between the time of its occurrence and the time of its recurrence. Every case of memory is thus a case of recurrence, but every case of recurrence is not a case of memory. In addition to the fact of recurrence there must be a recognition of recurrence, a recognition that something is being experienced *again*. The mere fact of recurrence is not difficult to understand. If the physiological ground or concomitant of an experience leaves as a trace of itself a specific modification of the brain structure— a "path," for example,—it is natural that when that particular part of the brain is re-excited, whether by stimulus or from within, we should re-experience the content. The really curious thing about memory is not the recurrence but the recognition of the recurrence; the consciousness of having experienced the *same* content at a more or less definitely localized moment of the past. Now we have seen that within the scope of a single specious present the amount of time which has elapsed since the occurrence of an event is measured by the ratio of its change-rate to the change-rate induced by it in the apperceptive system upon which it is "projected." This apperceptive system or apperception-mass is (when the attitude of attention is retrospective) nothing but the collective totality of contents which have been perceived since the event in question first appeared. And the change-rate of such a system is inversely proportional to its size. The specious present, then, increases with the increase of the apperception-mass, and hence the extent of the latter suffices to measure not only the duration of time from a point in the past up to the present, but conversely, the felt distance of that point from the present, *i. e.*, its date. Thus when the specious present is large, the date of its earlier events is felt to be remote, while, conversely, the earlier events of a small specious present are felt to be of recent date.

I believe that the date of a remembered object is measured in the same way. The specifically modified brain-cell, when

re-excited, excites in its turn the physiological traces of the experiences intervening between its first occurrence and the present moment, with the result that these intervening experiences vaguely and very schematically recur as an apperception-mass for the recurrence of the content which suggested them. And precisely as in the case of the specious present, the rate of change in this apperceptive system is, in proportion to its extent, small as compared with the change in the single content. The ratio of the latter to the former— $\frac{do}{ds}$ —measures and represents the date, *i. e.*, the period of time that appears to have elapsed since the event first occurred. But the remembered event is not only projected upon this large and schematic apperception-mass: it is, together with the latter, projected also upon the background of concrete experiences which fill the regular specious present of the moment. Thus it appears to us in *two* contexts or aspects—as a present event in our mental life and as an event of the past. The *againness* of the remembered event is due precisely to this duality of its apperceiving system.

The view that the date or temporal sign of a remembered happening is determined by the recurrence of the intervening experiences is borne out by the recognized facts of the memory consciousness. For, in the first place, we can recall the date of an event only when we have some memory of what has followed it. (By date I mean of course the *felt* date, not the date as abstractly known.) When the intervening experiences are numerous and interesting, the apperceptive system s, in which they schematically recur, is large, and the differential fraction $\frac{do}{ds}$, which measures the lapse of time since the occurrence of o, is also large. While, conversely, when few experiences have intervened between an event and our memory of it, the apperception-mass is small and the date appears more recent than it is.[1] Finally, when the activity of the perceptive faculty is completely suspended, as in trance or change of personality, and when consequently there are no intervening experiences out of which to make a memory apperception-mass, we find, as our view would lead us to expect, no appreciation of the period of time separating the remembered event from the moment at which it is remembered. In such cases, the events of months ago are regarded as those of yesterday.

In spite of the "concomitant variation" of the number of intervening experiences and the distance of the assigned date of the remembered event, it may still seem that we have not sufficiently explained the profound difference between memory of

[1] The rate of time-flow is, as we know, reversed in retrospect. Periods that felt brief when passing are remembered as long, and *vice versa*.

the remote past and immediate consciousness of the just past. In order to remove any such misgivings, let us sum up the points of difference already noted between the two kinds of consciousness of the past. In the first place, in immediate apprehension the apperception-mass is the concrete complex of experiences which having just occurred have not yet faded from consciousness. In memory, on the other hand, the principal apperception-mass is a vast, hazy, highly schematic group of contents, the merest skeleton of the experiences of the intervening period. And secondly, in addition to this contrast in the backgrounds upon which the two kinds of past are respectively projected, there is a single background in the case of immediate apprehension, while in memory there is a double background,—the schematic system just mentioned, and the concrete system of contents which makes up the specious present of the moment in which the act of memory takes place. And to this duality of apperceptive context is due, as we have seen, the feeling of "againness" or recurrence which is distinctive of the memory psychosis.

VIII.

Familiarity.

The feeling of familiarity, which is akin to memory, differs from recall proper in two ways. In the first place the event recognized as *familiar* is normally caused by an external stimulus, and hence possesses the vividness and arbitrary "givenness" characteristic of a sensation as distinguished from an image. And in the second place, the memory apperception-mass is so entirely subordinated to the specious present in which the familiar sensation occurs that it serves merely to give to the latter a fringe of pastness in which there is no appreciable indication of date. In short the feeling of familiarity is the apprehension of a present event with a fringe of pastness, or "recurringness," while in the memory judgment, the recognition of a more or less definitely dated past is primary, and is only secondarily qualified by a fringe of presentness due to the act of memory itself.

IX.

The Conditions under which Time Perception Originates.

When we inquire how a specious present begins or how it can apply to a single sensation, we are confronted by a peculiar difficulty. According to our view there must always be a pre-existing background of mental states upon which the content perceived as an object is projected. But in the first

moment of a psychosis, or in the case of a single sensation, this would seem impossible. And yet when we are awakened from a state of unconsciousness by a touch or a sound, we do not experience a sensation that is without duration. If anything, the time-form is more noticeable in the simple and approximately sensational first stage of an experience than in the later and more complex stages. The *meaning* of a sensory content is felt only when we feel its relation to other contents: but not so with its duration. As soon as it is perceived, or rather, in order that it may be perceived, we must be conscious of it as enduring. Our problem is to point out what it is that plays the part of apperception-mass in the case of a single sensation and at the first moment of an experience.

Let us begin by considering the case of a possible single sensation which has established itself in consciousness. Along with the sensation at any moment will be the consciousness of it as it was in the just preceding moments. Let o_4 represent the sensation at the fourth second of its existence, and o'_1, o'_2 and o'_3 symbolize the present image-traces of its earlier existence. It will be sensed in the light of these image-traces precisely as an incoming element in a complex pyschosis is apperceived by the other elements. The "apperceptive system" for a single sensation—if such were possible—would then be nothing but the image-traces of its just-past existence. And the specious present, $\frac{do}{ds}$, of such a sensation would be expressed (at the fourth second) by the symbol $\frac{do_4}{d\,(o'_1\, o'_2\, o'_3)}$. When the apperceptive system or subjective background, composed of the image-traces $o'_1\, o'_2\, o'_3$, became greater, or more prominent for consciousness than the present sensation o_4, the rate of change of the former would become less than that of the latter, *i. e.*, $\frac{do_4}{d\,(o'_1\, o'_2\, o'_3)} > 1$—and a specious present or perceptual time-form would be realized.

So much for the time-form of a single sensation after its origin. Can we make use of this account to explain the *simultaneous* origin of an experience and its time-form? Here as elsewhere, in explaining the origin of anything in consciousness, we are compelled to fall back upon those physiological processes which are the secondary, or it may be the proximate conditions of sensory states. When by the action of a specific stimulus an equally specific disturbance in the cortex is produced, the latter does not, we may suppose, die out the instant the stimulus ceases, but continues for a short time, as the violin string continues to vibrate after the bow has been withdrawn. If the same stimulus be applied again the resulting disturbance will be a complex of the pre-existing disturbance and the re-

newed effect of the stimulus. Now if the stimulus act in such a manner that the disturbance in the brain-center due to past action become greater than the effect of its present action, *there will be produced on the physiological level a perfect counterpart of the psychological conditions of the specious present.* That is, we shall have a "subjective" or pre-existing state of disturbance greater or more vigorous than the incoming impulse. The modification in the nerve center due to the just past will dominate that due to the present. If under these conditions a sensation arise, it will arise with a time form. Let us suppose a case in which a stimulus must act continuously for four seconds in order that a sensation may arise. We may represent the resulting disturbances of the central organ by $o(-_4)$, $o(-_3)$, $o(-_2)$, $o(-_1)$, and their traces by the same symbols primed. We may suppose that at the moment when the sensation o_1 arises, the physiological stimulus-traces $o'(-_4)$ $o'(-_3)$ and $o'(-_2)$ are together equal in intensity to the impulse $o(-_1)$, and that when increased by the trace of the latter the pre-existing state will dominate the incoming state,—which, if it be felt at all, will be felt in a time-form or specious present represented by

$$\frac{do}{d\,(o^1_{(-4)}\ o^1_{(-3)}\ o^1_{(-2)}\ o^1_{(-1)})} > 1.$$

Now inasmuch as every sensation which we can imagine ourselves to experience has a time-form, we must conclude that in defining the conditions of the origin of a specious present we have at the same time defined one of the essential conditions for the origin of conscious states in general.

The several steps in our argument may be summed up as follows:

I. The central fact of time-consciousness is the "specious present"—the present consciousness of what is no longer present. Without this apparent projection of past moments upon each present moment, there could be no consciousness of either duration or succession.

II. Time is the form of change; and change, like every magnitude, is relative to its unit of measure. The amount of perceived change in an object is relative to the actual concomitant change in the perceiving subject.

III. Regarding the 'object' as simply the objective aspect, and the 'subject' as the subjective aspect of a psychosis, we may hold that each incoming objective content is apperceived by or projected upon the entire system of pre-existing states. The psychosis considered only in its collective unity is subjective; considered only in its distributive plurality, it is an aggregate of objects; considered (as it must be) under both aspects simultaneously, it is a "subject-perceiving-objects." As the apperceptive system is larger than any of its objective contents,

the constant stream of increments and decrements produces in the latter changes that are proportionately greater or faster than the change that they in turn cause in the system as a whole. *This difference in the rates of an objective change and the subjective change which reflects it, measures the amount of perceived change and perceived time.* And though $\triangle s$ and $\triangle o$ (the actual subjective and objective changes respectively produced in s and o by the environmental increment of content) would approach zero as $\triangle t$ approached zero, yet the ratio $\frac{\triangle o}{\triangle s}$ would approach or maintain the finite value $\frac{do}{ds}$. And this for the same reason that the velocity of a moving body, if uniform, maintains, and if accelerated, approaches, a finite value, when the space through which and the time in which it moves approach zero. By assuming the amount of perceived time to be a continuously present *differential rate*, it becomes possible to reconcile the "specious," "vicarious," or "perceptual" presence of an extended segment of the past with the real or unextended present. Finite positive portions of time and change are perceived as present when $\frac{s}{o}$, and consequently, $\frac{do}{ds}$ are greater than unity.

IV. Time is felt as *enduring* when $\frac{do}{ds}$ is increasing in value; *i. e.*, when $\frac{d^2o}{ds^2}$ is positive; and conversely, when the second derivative is negative, the events are felt as escaping or *passing*.

V. As the positive and negative *signs* of the second derivative respectively denote the enduring and the passing phases of perceptual time, so the *value* of the same ratio measures the perceived *rate* of time. Time "flies" or "drags" according as $\frac{d^2o}{ds^2}$ is large or small. And as in this case the first and second derivatives vary inversely—*i. e.*, the longer the time-span or specious present the slower its rate of change—it follows that the events that are near to the present appear to *acquire pastness* more rapidly than those further past. (This I take to be the explanation of the apparent conformity of time-perception to Weber's Law.)

VI. A specious present tends constantly to increase, but this naturally cumulative tendency brings about automatically in every finite consciousness a counter-tendency; for in so far as the contents of a psychosis are not perceived in the light of a single plan, the degree to which each is attended to is inversely proportional to the number of all; and there comes a time when the content of longest duration can no longer hold the attention.[1]

[1]There are two fundamentally distinct features of a content that attract the attention—its intrinsic intensity and its extrinsic bearing upon other contents. The latter is the conceptual or unifying factor and is usually possessed in the greatest degree by the content of longest duration, because it is the center in relation to which the rest of the

That activity on being thus torn from its attachment to the keystone of the system, by the dissipative force of numbers, is for the moment open to the highest bidder, which is usually the most recent sensation; and about this as a center a new present is built up. This alternation of slow accumulation and rapid crumbling, in the time-span, throws light alike upon the rhythmic character of consciousness and upon the pleasurableness of objective rhythm.

VII. Memory differs from the immediate consciousness of the past in two ways. In the first place the remembered event (*i. e.*, the re-excited trace of the past experience) is projected upon a schematic and hazy apperception-mass, composed of traces of a few of the more important intervening experiences. This makes possible the felt date, while it enables us also to account for the indefiniteness and inaccuracy of it. In the second place the remembered event is less noticeably projected upon the concrete apperception-mass of the specious present in which the act of memory takes place. It is thus seen in two contexts, and to this duality of background its "againness" may be ascribed.

VIII. The feeling of familiarity differs from memory proper, in that the trace of the past event is excited sensorily from without, and is projected primarily upon the specious present of the moment, and only secondarily upon the background of intervening experiences. It is an experience of the present with a fringe of the past, while memory is an experience of the past with a fringe of the present.

IX. The manner in which a time-form can apply to a single sensation, and the manner which it can originate at all, present peculiar difficulties. Confining our attention to a possible single sensation, we find that in such a case the apperceiving background for the sensation, after its origin, would be composed of the image-traces of the sensation itself. The formula for the specious present becomes $\frac{do}{do'}$ where o^1 represents the pre-existing subjective traces due to the presence of the sensation in the preceding moments. But these quantitative symbols may with equal or, as some psychologists would hold, with greater propriety be used to denote the physiological concomitants of sensation. And the symbol $\frac{do}{do'}$, when taken physiologically, may represent a condition just *antecedent* to perception. From this point of view we can understand how it is that the time-form of a psychosis arises simultaneously with its contents.

---

system is interpreted. On the other hand the intensive or sensory factor is peculiarly the property of the sensation of the moment—and it is the gradual but inevitable gain of the latter or centrifugal over the former or centripetal factor, that brings about the periodic dissolution of each psychosis.

AUDITORY TESTS.[1]

BENJAMIN RICHARD ANDREWS, A. B., A. M.

TABLE OF CONTENTS.

		PAGE.
Introduction.		14
Chapter I.	Mental Tests and Their Relation to Psychology and Pathology	15
Chapter II.	Literature of Mental Tests	20
Chapter III.	Tests of Acuity of Hearing	26
	A. Speech Tests of Acuity	26
	B. Tests with Mechanical Sounds	42
	I. Tests of Acuity with Watch	42
	Acoumeter	44
	Audiometer and Electrical Devices	45
	Forks	46
	II. Tests of Tonal Limits	46
	Lower Limit	46
	Upper Limit	48
	C. Appendix to Chapter III.	49

INTRODUCTION.

The two great functions of the ear are these: it furnishes the most common means of intercommunication, that of spoken speech; and it opens to mankind one whole department of the fine arts, that of music.

Efficient hearing is indispensable if the individual is to do his full part in the world's work, or to enjoy his share of its æsthetic pleasures. Normal functioning of the auditory organs is, therefore, a matter of importance to the individual. This paper discusses tests which may be applied to an individual to determine his auditory capacity as normal or abnormal. Such tests form part of what are commonly called "mental tests," *i. e.*, tests which have to do with mental factors. Our subject has, therefore, a certain relation to psychology. Moreover, since certain auditory tests are used to diagnose aural disease, it has also a relation to pathology.

The first chapter contains an examination of the concept of "mental tests," and of the relation of such tests to psychology and pathology. In the second chapter, the literature of mental tests is briefly reviewed. These two chapters serve as a pref-

[1] From the Psychological Laboratory of Cornell University.

ace to the description of the various auditory tests, which forms the main problem of the paper. These tests are treated in three divisions, forming chapters three, four and five: Chapter III: Tests of auditory acuity (a) with spoken words and (b) with mechanical sounds; Chapter IV: Tests of musical capacity; Chapter V: Tests used in diagnosis of aural disease. The present paper includes Chapters I-III; Chapters IV and V will follow in a subsequent article.[1]

CHAPTER I.

Mental Tests and their Relation to Psychology and Pathology.

The term "test" implies the existence, or possibility, of an average quantitative expression for a physical or mental characteristic or capacity, to which as a standard or norm the measurements of given individuals may be referred. The standard or norm value is taken as the average or mean of the measurements of many individuals with regard to the characteristic concerned. This concept lies at the basis of anthropometry.

This science has confined itself thus far almost entirely to physical measurements of the human body. Anthropometrists have compiled tables which show the normal height, weight, strength, lung capacity, length of arm, and other purely physical characteristics, of the typical individual, in various classes arranged according to age, sex, race, occupation and other differentiae. The normal proportions between various measurements, as those of girth and height, have also been recorded. Such norms are ideal values secured by taking the average or mean value of the particular measurements in a large number of individuals of a definite class; they represent the development which the normal individual within the class is most likely to possess. When norms are once established, the measurements of given individuals can be referred to them, and the individuals classed as normal, or as sub-normal or super-normal in a certain degree, with respect to the particular characteristics. The results of physical anthropometry are valuable, first of all, as scientific facts: like the facts of astronomy, they have a value in and of themselves. They find a considerable practical application as well. Successive testing of the same individual measures the influence of exercise, employment, or habits

[1]The writer wishes to express his obligation to Professor I. Madison Bentley, of the Department of Psychology, Cornell University, for putting at his disposal the contents of an unpublished paper on tests of audition, which proved very valuable in giving a starting point for the writer's own investigations.

of life, on his physical organism. Anthropometrical examinations guard the entrance to certain occupations, as the army and police, and in some instances admission into public schools. Such measurements have been adapted by A. Bertillon to furnish an accurate system of identification of criminals. They have recently been introduced into public schools to detect abnormal bodily development and make physical training effective; and their results are throwing light on other pedagogical problems. In all such measurements, there is assumed an average typical development for members of each class, and individual variations from this norm are regarded as of significance. In these physical measurements, which comprise thus far the great bulk of anthropometrical investigations, there is coming to be general agreement as to the dimensions and proportions which shall be noted, their normal values, and, to some extent, the significance of deviations from these values.

Anthropometry has shown also a tendency to use tests involving mental factors and, indeed, to measure mental characteristics themselves. Francis Galton included certain mental measurements in the anthropometrical investigations which he suggested in 1882[1] and afterward carried out. In 1890 the matter was brought to the attention of psychologists by Cattell's article in *Mind*[2] on "Mental Tests and Measurements." He had already applied measurements to typical mental characteristics, and in this article advocated the adoption and wide use by psychologists of tests of vision, audition, reaction time, and other mental phenomena. Other investigators took up the matter, and a few years later, in 1895, the American Psychological Association appointed a committee to select a series of tests through which there could be established "the normal capacity of simple and typical sensory, motor and intellectual endowment in the average individual and in groups; the distribution of these capacities, their correlation, development, and relation to daily activities."[3] The tests were to be used in psychological laboratories upon large numbers of college students and other individuals, and in this way data secured "for the study of general, typical, characteristic endowment, much as the anthropometrist desires first to establish standards of the principal physical measurements and proportions."[4] Such tests were instituted at Columbia, Wisconsin, and some other universities. A programme closely related to that of "mental tests" is the individual psychology proposed in France

[1] *Fortnightly Review*, xxxvii, 1882, 332.
[2] *Mind*, XV, 1890, 373.
[3] J. Jastrow (member of committee): *Psychological Review*, V, 1898, 172.
[4] *Ibid.*

by Binet and Henri and in Germany by Kraepelin and Oehrn. Its problem, as stated by Binet and Henri is (1) the study of the variable properties of psychical processes,—how and to what extent processes vary from individual to individual; and (2) the discovery of the interrelations of these variable properties in the individual mind. Without reviewing the literature of either mental tests or individual psychology at this point, we may say that each proposes to measure such mental functions as memory, association, reaction-time, imagination, audition, vision, attention, etc., in large numbers of individuals; to determine norms as physical anthropometry has done; and to discover and interpret individual variations from them. Each, moreover, expects to contribute to the science of psychology. It will, therefore, be desirable to determine at once how mental tests are related to psychology and whether they come within its scope as a science.

Now modern psychology may take three more or less clearly defined attitudes toward its subject matter, mental experiences. (1) It may regard mental stuff statically, as structure, disregarding its reference in the environment and its significance to the individual. It then finds its immediate problem in the analysis of consciousness into structural elements, the sensations and affections, and in the statement of the laws governing their interconnection in the structural complexes found in consciousness. (2) It may regard the processes dynamically, as functions of the psychophysical organism, having meaning to the organism in its relations to the outside world. The problem, as before, is, first, analysis, but now into elementary functions; and secondly, a statement of the laws governing their interaction. (3) It may regard mind as in process of development and trace its progress genetically in the race or in the individual. Since this genetic attitude may be either structural or functional, the primary psychological standpoints reduce to two. To which of these are mental tests related? In general, not to structural psychology, we may say, since one tests the individual in some mental characteristic regarded in its significance to him and without regard to its structural aspects, the sensations and affections which make it up. The answer is not so easy regarding the relation of mental tests and functional psychology. They work upon common material, namely, mental functions. They may be concerned with the same function. Yet, from the standpoint of scientific methodology, with which the decision must finally rest, mental tests and functional psychology lie outside each other, and mental tests, as tests, are not a part of the science of pure psychology. Psychology approaches its problem with trained observers, delicate apparatus, careful procedure designed to check errors, rigid experimental conditions,

and with the single purpose of seeing and describing the phenomenon as it is. A mental test of the same function involves the examination of untrained subjects, under limitations of time which make refined measurement and precise procedure impossible, and with the subjective conditions of tranquility, attention, and knowledge of the problem varying widely from observer to observer. Norms are based upon such measurements, and the individual's variation from them is interpreted for some practical end, perhaps medical diagnosis, or a characterization of the individual as of normal capacity in a certain direction. Psychology, on the other hand, does not recognize individuals as such. Just as it is not a fact in the science of theoretical physics that a certain engine produces a certain amount of power per fuel unit, so it is not a problem for psychology to evaluate John Smith's memory or attention. The former problem falls to the province of engineering, an application of pure physics; the latter, a typical mental test, is just as certainly an application of psychology, and is not to be regarded as within its limits as a theoretical science. Further, psychology of function, as a science, has no room for the practical, extra-scientific ends involved in mental tests; it measures mental processes simply to make possible a complete statement of the functional elements of mind and their interaction.

This distinction, which puts mental tests outside psychology, does not exclude the possibility of a scientific individual psychology, which, as Henri and Binet propose, shall describe the variations of mental processes from mind to mind. The objection is simply against the attempt to found such a psychology on the rapid, necessarily inaccurate, and wholesale testing of miscellaneous individuals. The painstaking determinations of trained psychologists, working with refined methods and selected observers, may alone be expected to build up a trustworthy body of facts for a psychology of the individual. The progress of child psychology is illuminating and admonitory in this connection. The established facts regarding the child's mind have been secured by a half-dozen trained investigators through long-continued observations of children; while the mass of facts secured by the questionary and the untrained observations of parents and teachers are largely either unreliable or outside of psychological science. So, doubtless, the development of functional psychology, whether individual or general, will depend on the investigations of experienced psychologists, working with practised observers, and not on the statistical treatment of measurements made on large numbers of individuals taken at random. Mental tests, as tests, therefore, are not to be regarded as part of psychology, neither should direct contributions to psychology be expected from them. Psychology, on the other

hand, can do a great deal for mental tests. In the first place, it must guide in selecting the tests to be employed; and, secondly, it must furnish appliances and methods of procedure, to be adapted as the grosser work of the tests may require.

In concluding this discussion, then, we may say that mental tests, as tests, are more akin to physical anthropometry than to psychology. They are to be sharply distinguished from the measurement of mental functions in psychology, both because of the differences in refinement of method, and particularly because they are applied to masses of individuals to determine norms, individual variations from which are interpreted with regard to some practical end. Mental tests may be considered as an application of psychology, and in so far they are related to it. In the methodology of science, it seems proper to term them, as has been suggested, mental anthropometry.

There remains the further question of the relation of mental tests to pathology, in particular to medical diagnosis. In discussing the applications of physical anthropometry, we mentioned that it has been found useful, as in gymnasiums, in detecting poor physical development, in indicating the exercise or other treatment needed, and in measuring the response of the physical organism to such treatment. Certain tests involving mental factors are used in a very similar way to determine the presence of diseased or abnormal conditions, as the tests of vision, audition, and the rough tests of intelligence to determine insanity. The oculist uses a series of tests whose results indicate primarily the normal or abnormal condition of the visual organs, and, further, if the eye be abnormal or diseased, reveal the nature of the disease or abnormality. Similiarly, the aurist employs tests which indicate whether hearing is sound and normal, and certain other tests which, in case of diseased or abnormal conditions of the ear, reveal the origin, nature, and extent of the disturbance. In such tests, at least so far as diagnosis of disease is involved, we are in the field of pathology. Because of their significance in identifying diseased conditions, they may be called diagnostic tests. While such determinations are primarily matters of pathological practice, yet they involve mental factors, and so become of interest not only for their own sake, but for their relation to psychology. In any case, they should be included, to make our systematic presentation complete. In our treatment of auditory tests, we shall accordingly discuss, first, tests measuring the normal condition of the auditory organ and variations from the normal; and, secondly, certain tests useful to otology, the special branch of pathological science treating of aural disease.

CHAPTER II.

LITERATURE.

This sketch of the literature of mental and auditory tests will not attempt an exhaustive review of all that has been written on the subject. The general bearings of both mental tests and individual psychology, and their literature down to 1898, have been reviewed by Miss Sharp in her article on Individual Psychology;[1] and the literature of auditory tests up to ten years ago, has been summarized by Chrisman.[2] Further, as we come presently to consider various tests of audition, we shall be in constant reference to the literature. For these reasons, a detailed examination of all the literature would be superfluous. At the same time, a brief review of the principal statements of mental tests will be valuable in giving perspective and in relating the present work to that of previous writers, particularly American.

Mental tests, as already stated, go back to Francis Galton. His article on "The Anthropometric Laboratory" in 1882,[3] urged the establishment of laboratories where measurements of purely physical characteristics and of certain mental characteristics as well, could be made. The tests which he suggested include, in addition to purely physical measurements, the following: strength by dynamometer, muscular agility, co-ordination of eye and muscle, sight, hearing, touch, muscular sense, reaction-time, after-images, and memory. Galton's tests of hearing embraced "keenness of hearing, and the appreciation of different degrees of loudness and of different tones." It is to be noted that Galton's standpoint, in suggesting such tests, mental and physical, is that of the anthropometrist. He called attention to the obligation of mental tests to psychology, a debt no one disputes, in his remark that from the work of Fechner and Delboeuf and similar investigators suitable instruments could be selected for "an anthropometric laboratory." He said nothing of possible contributions from mental tests to psychology; this idea was developed, as we shall see presently, by later writers whose interests were in psychology, not in anthropometry. It is important to note that mental tests were first suggested by Galton as anthropometric measurements, quite on a par with physical measurements, without a thought that they would be anything more. This may be fairly inferred from the reasons which he gives for performing such

[1] S. E. Sharp: *American Journal of Psychology*, 1899, 329-348. See also her conclusions, pp. 388-391.

[2] Oscar Chrisman: *Pedagogical Seminary*, ii., 1892-3, 397.

[3] *Fortnightly Review*, xxxvii, 332; reprinted in *Popular Science Monthly*, xxi, 53.

tests.[1] It would be interesting to follow Galton's development of the idea of mental tests further; but it must suffice to call attention to his book of 1883, "Inquiries into Human Faculty,"[2] and to the establishment of a temporary anthropometric laboratory in 1884 in connection with the International Health Exhibition at South Kensington, London. At this laboratory, various measurements, some purely physical, some involving mental factors (*e. g.*, the highest audible tone), were made upon over 9,000 persons.[3] The anthropometric laboratory was maintained at South Kensington until about 1895, and then removed to Oxford.

With Galton, mental tests were simply a branch of anthropometry. Their attention at the hands of psychologists seems first to have been due to Cattell. In 1885, and subsequently, he published the first contributions to individual psychology,[4] and in 1890, his article in *Mind*[5] brought the subject of mental tests to the attention of psychologists. Cattell's article was at once an outline of tests already performed, and a suggested programme of tests for general use in the psychological laboratories then forming in American universities. Cattell's tests all involved mental factors, some, as dynamometer pressure, including as well prominent physical factors; others, as memory, being purely mental. It is important to note that his purpose in performing the tests was primarily psychological, to contribute to the science of psychology. He says: "psychology cannot attain the certainty or exactness of the physical sciences, unless it rests on a foundation of experiment and measurement. A step in this direction could be made by applying a series of mental tests and measurements to a large number of individuals. The results would be of considerable scientific value in discovering the constancy of mental processes, their interdependence, and their variation under definite circumstances." While recognizing the practical possibilities of tests, Cattell here apparently puts considerable faith as well in their "scientific value" for psychology. This was in 1890. We may fol-

[1] Galton's reasons are: (1) biographical interest; (2) medical interest to the person tested; (3) information for descendants as to hereditary dangers and vital probabilities; (4) material for investigations into life histories.

[2] Now, unfortunately, out of print.

[3] *Anthropological Institute Journal*, xiv, 1884-5, 205, 275; *ibid.*, xvi, 2. For other references to Galton's work, *cf. Science*, v, 294, and viii, 374; *Nature*, xxxi, 223. *Cf.* also *Lippincott's Mag.*, xlv, 1890, 236, *Why We Measure Mankind*, by Galton.

[4] *Cf.* foot-note, *Psych. Rev.* iii, 1896, 618. Cattell here says that the introduction of individual psychology was delayed because Wundt "was not favorable to it."

[5] J. McKeen Cattell: *Mind*, xv, 1890, 373, *Mental Tests and Measurements.*

low this trend toward a psychological evaluation of mental tests further.

The idea of contributing to psychology by wide-spread testing and measurement seemed to secure general favor. Jastrow, in 1891, published the results of some tests on the community and association of ideas in men and women.[1] In 1892, there were published the results of a series of anthropometrical and psychological tests carried out under Jastrow's direction on students at the University of Wisconsin.[2] The next year, Jastrow conducted a psychological laboratory at the Columbian Exposition in Chicago, in part similar to Galton's anthropometric laboratory, but giving more attention to strictly psychological tests.[3] In the same year (1893), there were two statements of mental tests, by Scripture[4] and Cattell[5] respectively, especially in their relation to school children. The practical value of tests was urged by both, and both seemed impressed with "the scientific value of such tests on large numbers of persons," as Scripture broadly puts it. Cattell, indeed, states explicitly the position with which the writer has taken issue. In urging teachers to perform tests upon school children, he says, "those who wish to contribute to the advancement of psychology will find this a convenient opening." This is a rather radical statement of the relation of tests to psychology, but it seems to represent the prevalent feeling of the time that in mental tests and statistical methods psychology had a new organon. At this juncture, when the proposed psychological laboratory at the Columbian Exposition was under consideration, Titchener published an article on the relations of psychology and mental tests in which there are stated, explicitly

[1] Joseph Jastrow: *A Study in Mental Statistics, New Review*, v, 1891, 559; *A Statistical Study of Memory and Association, Educational Review*, ii, 1895, 442. Another study by Jastrow was published in 1894: *Community and Association of Ideas: A Statistical Study, Psych. Rev.*, i, 1894, 152. This last test was repeated at Wellesley College with different results, under the direction of M. W. Calkins: *Psych. Rev.*, ii, 1895, 363. A criticism of the latter by Jastrow followed: *ibid.*, iii, 1896, 68; to which Miss Calkins replied, *ibid.*, 426. Jastrow made another reply: *ibid.*, 430. The Wisconsin and Wellesley tests are criticised by Amy Tanner: *ibid.*, 548.

[2] *Am. Jour. Psych.*, iv, 1892, 420. The tests were performed in the fall of 1890.

[3] "Experimental Psychology at the World's Fair," an outline of plans given in an address by Jastrow before the American Psychological Association, Philadelphia, December 27, 1892. / *Cf. Official Catalogue of the World's Columbian Exposition*, pt. 12, 50. For an analysis of the tests, *cf. L'Année psychologique*, i, 522. No results, as far as the writer has been able to learn, have been published.

[4] E. W. Scripture: *Tests of School Children, Educ. Rev.*, v, 1893, 52.

[5] J. McK. Cattell: *Tests of Senses and Faculties, Educ. Rev.*, v, 1893, 257.

or implicitly, many of the distinctions urged in our preceding chapter.[1] He raises the question pointedly: "What constitutes the difference between a psychological and an anthropometrical experiment?" and finds very valid differentiation. The psychological experiment assumes, on the part of the observer, "practice in introspection, in attentional concentration, and in control of the particular apparatus employed. The anthropometric measurement requires at most only enough practice to carry out instructions. The value of the former is as much qualitative (for analysis and description) as quantitative; the value of the latter is solely quantitative." Speaking further of the line between experimental psychology and anthropometry, he says, "It is as unfair to rob anthropometry of her experiments and dub them psychological, as it is to psychology to conceal the difficulty of her own experimentation by substituting the simpler anthropometrical for it." Though this article was written with the simple aim of making suggestions as to the psychological exhibit at Chicago, it had in it a wider significance, in implicitly drawing the line of demarcation between the exact introspective experimentation with trained observers which lies within psychology, and approximating measurements of mental capacity, made upon miscellaneous individuals, which lie outside psychology. If the validity of the distinction suggested in Titchener's article had been recognized, mental tests would not have been welcomed, as they were, as a new and promising psychological method. The distinction was not recognized, however, and mental tests for a time had their welcome in psychology. We may follow their fortunes further.

Gilbert, during 1893-4, made tests of the physical and mental development of New Haven (Conn.) school children.[2] These tests, he says, were "to aid in the analysis of mental phenomena,"—a purely psychological aim, surely. In December, 1895, the American Psychological Association appointed a committee, composed of Cattell, Baldwin, Jastrow, Sanford and Witmer, "to consider the feasibility of co-operation among the various psychological laboratories in the collection of mental and physical statistics."[3] Cattell and Farrand, in 1896, published a report of physical and mental tests on students of Columbia University.[5] In this article, they state the purpose of

[1] E. B. Titchener; *Anthropometry and Experimental Psychology, Philos. Rev.*, ii, 1893, 187.

[2] J. A. Gilbert: *Studies from Yale Psych. Lab.*, (Nov. 1894) II, 40-100. His mental tests measure muscle-sense, color-perception, suggestion, motor ability, fatigue, reaction-time, discrimination-time and i me memory.

[3] *Psych. Rev.*, iii, 1896, 122.

[4] *Ibid.*, 618.

such tests to be "the study of the development and correlation of mental and physical traits. . . . . The tests have two chief ends, the one genetic, the other quantitative. We wish to study growth as dependent on environment and heredity, and the correlation of traits from the point of view of exact science." The authors seem to have accepted something of the distinction urged by Titchener three years before, and to have in mind the limited psychological bearing of mental tests, for they refer to them as "anthropometric work." The reasons urged for the performance of tests are largely practical rather than scientific: they say that the tests are of interest and value to the individual; may make the psychological laboratory of value to the community in school tests, medical practice, etc.; and give advanced students practice in research. On the other hand, they find certain psychological ends for mental tests, since they are to study (1) the interrelation of the traits defined and measured, and (2) the development of the individual and the race. The first of these aims falls in with individual psychology, the second with genetic; and so far the authors seem to hold to the psychological possibilities of mental tests.

The committee of the association presented its preliminary report in December, 1896.[1] A second report, followed by discussion, was made at the meeting a year later.[2] In the former report, the purpose of mental tests is stated briefly; "to reveal individual differences and development." On the latter occasion, Jastrow gave a paper on "Popular Tests of Mental Capacity," which was supplemented by statements from Baldwin and Cattell. Jastrow was at pains to differentiate "the careful and ingenious analysis on the part of well trained and scientifically self-observant experimentalists," to which he says the main equipment of a psychological laboratory may wisely be devoted; and investigations "to establish the normal capacity of simple and typical sensory, motor, and intellectual endowments," in the average individual or in selected groups, and the distribution, development and correlation of such powers,—in short, tests or "mental anthropometry." Though in this statement, Jastrow distinctly separates tests and psychology, it is doubtful whether the distinction has been clearly recognized in psychological thought at large. Jastrow himself in his latest statement says:[3] "Call it mental anthropometry if you will, but do not disregard the valuable contributions to other divisions of psychology and to the general conceptions of this science which such investigation has the possibility, and in my opinion the probability, of contributing." He thus

[1] *Psych. Rev.*, iv, 1897, 132-138.
[2] *Ibid.*, v, 1898, 145, 146 and 172-179.
[3] *Ibid.*, viii, 1901, 14.

maintains the possibility of contributions to psychology from mental tests, as does Cattell in the last statement quoted from him. On the other hand, the fact that seven years have passed since the committee suggested its tests, and that psychologists have not made any considerable contributions in this field indicates that, whether consciously or not, the distinction between mental anthropometry and psychology has come to be accepted. Some tests have indeed been performed.[1] One of the investigators, Seashore, seems clear with regard to the distinction just referred to: "most of the development of a problem," he says, "can be carried on to the greatest advantage by the ordinary method with trained observers; it is mainly when we wish to determine the nature of the naïve experience or the uniformity of a certain tendency that we find it profitable to appeal to the statistical method."[2]

There are indications, indeed, that the practical purposes in mental tests are coming to be recognized as the real and only reason for their performance; and that the hope of making them contribute to psychology is being given up. The practical end has been recognized from the first by Galton, Cattell, Jastrow, and others, as a large reason for performing tests. It seems fair to say that the present attitude toward mental tests, in so far as there is such an attitude, emphasizes the practical and minimizes the psychological aspect. The fact that psychologists have gone into public schools with tests and urged practical benefits from their performance; the fact that, despite "possible and probable contributions" to psychology, one does not actually find that such contributions have been made;[3] the fact, finally, that the public schools themselves are instituting mental tests for practical purposes,[4] indicate that mental tests are finding their place as a form of anthropometry carried out on large bodies of individuals for practical ends, and not as

[1] J. A. Gilbert: *Researches on School Children and College Students, Univ. of Iowa Studies in Psych.*, i, 1897, 1; C. E. Seashore, *Some Psychological Statistics, ibid.*, ii, 1899, 1; E. A. Kirkpatrick, *Individual Tests of School Children, Psych. Rev.*, vii, 1900, 274.

[2] *Univ. of Iowa Studies in Psych.*, ii, 1899, 3.

[3] One of the latest contributions to mental tests is C. Wissler's *Correlation of Mental Tests and Physical Tests, Columbia University Dissertations, 1901.* Wissler compiles the tests made for several years at Columbia by Cattell and others and concludes that (1) the laboratory mental tests show little inter-correlation; (2) the physical tests show a general tendency to correlate among themselves but only to a slight degree with the mental tests.

[4] The most striking example of this is the Department of Child Study and Pedagogic Investigation, of the Chicago Public Schools. *Cf.* its reports, 1898-99 ff. *Cf.* also A. MacDonald: *Report of U. S. Commissioner of Education*, 1897-8, 989 ff. and 1281 ff. Tests on a less extensive scale have been made in other schools.

a branch or method of pure psychology. If the discussion in this paper helps to make the distinction between psychology and mental anthropometry more explicit, and calls attention to the emphatically practical purposes of the latter, one of its aims will have been accomplished.

The present paper is an attempt to describe, criticise, and arrange tests of audition, which will admit of application by unskilled investigators. It recognizes the practical purposes of tests and formulates testing methods for extra-psychological ends only. It is not a contribution to psychological science, but to an application of psychology, mental anthropometry. This may itself in time become a scientific discipline; and this paper is offered as a modest contribution toward that end. As already stated, the auditory tests which will be considered are of two classes, anthropometrical and diagnostic. The former fall into two divisions: those of general auditory acuity, and those of musical capacity. These will be treated in Chapter III and Chapter IV respectively. Diagnostic tests occupy the concluding chapter, Chapter V.

CHAPTER III.

Tests of Acuity of Hearing.

This chapter falls into two parts: (a) tests which employ speech in examining acuity of hearing; and (b) tests of acuity with mechanically produced tones and noises.

A. Speech Tests of Acuity of Hearing.

The most important function of the human ear is the hearing of conversational speech. No examination of the auditory organ is entirely adequate unless it determines the practical efficiency of this capacity. Impaired acuity of hearing ordinarily first manifests itself by inability to hear speech distinctly and easily. Further, in some aural diseases, the ear no longer hears certain vocal elements, *e. g.*, the low, deep sounds of speech, or words containing them. Tests with spoken words may be employed to disclose either of these disturbances. Speech tests for diagnostic purposes, *i. e.*, to determine those vocal elements which have become imperceptible through disease and to interpret their loss as indicative of certain aural diseases, will be considered in a later section, under diagnostic tests. Here we are concerned only with tests employing speech to measure the degree of general auditory acuity. In such tests, words are commonly spoken at various distances from an individual and the maximal range is determined at which he hears the words correctly.

The reasons for employing speech tests may be briefly stated:

(1) They measure directly and unequivocally the most important function of the ear,—the hearing of conversational speech.

(2) Other tests do not measure this functional capacity. The results of tests with the watch, acoumeter, audiometer, tuning forks, and other mechanical sources of sound may not be considered an unequivocal measure of the individual's ability to hear speech.[1] Thus, one may hear the watch at a considerable range, and yet be relatively deaf for speech, or conversely. The ratio between the range for speech and that for a single tone, or noise, is not necessarily constant, even in the same individual. The reason for this is evident. Hearing a watch or acoumeter click, or a single tone, involves the perception of a limited number of auditory qualities, and, assuming the Helmholtz theory of audition to be correct, the normal functioning of a restricted and constant part of the basilar membrane. Perception of speech, on the other hand, requires the hearing of a large number of tonal qualities and noises of high, low and medium pitch, and, accordingly, unimpaired physiological functioning over a wide stretch of the basilar membrane. Further, defects in other parts of the auditory organ, as in the drum, or the chain of ossicles, might accommodate the transference of sound of one pitch, say that of the watch, and prevent the passage of certain sounds in the wider range of speech. For these reasons, speech tests alone are entirely reliable in determining the functioning condition of the ear for practical usefulness in life. Accordingly, while conditions will often dictate the use of the mechanical tests of hearing, explained in the next section, speech tests should, where possible, be given the preference.

The problem of arranging speech tests is as follows. The syllables, words, or sentences to be used in the tests must be chosen; the conditions under which the tests are to be applied must be stated; a large number of individuals must be tested and the normal range for speech perception determined, so that the individual's range, and hence his acuity of hearing, can be interpreted in terms of the norm. This problem is more complicated than it seems at first to be.

The selection of speech tests is difficult, first of all, because of the great variety and complexity of sounds involved in human speech. The elements of spoken speech vary in pitch, intensity, complexity, and clang-tint or 'quality.' Some elements, the vowels, are approximately pure clangs with faint

[1] *Diseases of the Ear*, D. B. St. J. Rosa, New York, 1885, 49-50. *Diagnosis and Treatment of Ear Diseases*, Buch, New York, 1880, 17. *Ohrenheilkunde*, A. Politzer, Stuttgart, 1893, 115, 119. *Schuluntersuchungen ü. d. kindliche Gehörorgan*, F. Bezold, Wiesbaden, 1885, 9.

traces of noise-elements; others, the consonants, are noises with more or less slight admixture of tonal elements.[1] As spoken in words and sentences, speech presents further complications due to accent, emphasis and inflection. The classical investigation of speech in its relation to audition is that of Oscar Wolf.[2] Its most important feature is a statement (*a*) of the pitches of the elements of speech and (*b*) of their intensities. (*a*) *The pitch of speech-elements.* Wolf accepted the pitch determinations for vowels made by Helmholtz, and made similar determinations for all consonants having vowel elements. The results are expressed in musical symbols or vibration rates, *e. g.*, the vowel "a" has a pitch of b^2 or 896 vibrations; the consonant "b," a pitch of e^1 or 320 vibrations. The various vocal elements extend in all over eight octaves, from the lowest sound, lingual-R which contains a tone of 16 vibrations, to the highest, the S-sound, with a tone of c^6 or 4,032 vibrations. In Wolf's classification, the consonants h, l, m, n and w are assigned no pitch, as they are in themselves devoid of characteristic tone, and assume that of the vowel to which they are associated. (*b*) *The intensity of speech-elements.* Wolf measured this by the comparative distances over which the various elements of speech carry and are still audible.[3] The vowel "a" (as in father) has the greatest range, 252 m.; the aspirate "h," the shortest range, 8.2 m. The other sounds fall between these limits, the vowels having much longer ranges than the consonants. The vowels, indeed, may be distinguished at distances at which the consonants are all inaudible. The consonants, further, vary very widely in their relative ranges. These determinations of Wolf, which furnish a complete inventory of the pitches and intensities of the elements of speech, have made possible its accurate application in auditory tests. Later investigation will doubtless set aside parts of Wolf's work, particularly the pitches assigned to the various vocal elements. L. Hermann,[4] for example, has analyzed the vowels phono-

[1] W. Wundt: *Grundzüge d. phys. Psych.*, ii, 1893, 49.

[2] *Sprache u. Ohr*, O. Wolf, Braunschweig, 1871; and various articles in otological journals, particularly, *Arch. f. Augen- u. Ohrenheilkunde*," III, Abth. 2, 35; and IV, Abth. 1, 125; and *Zeits. f. Ohrenheilkunde*, XX.

[3] C. Blake made a mechanical analysis of the relative intensity of consonants (*Zeitsch. f. Ohrenheilk.*, XI, 1 Heft). His results do not agree very well with those of Wolf. The latter's results are certainly to be followed in making speech tests, since he measured intensity by the relative distance at which the speech elements are heard by the human ear.

[4] *Pflüger's Archiv*, xlviii, 371 and liii, 1. Summary in Wundt, *op. cit.*, II, 49, 51, where a table compares Helmholtz' and Hermann's results. See also Helmholtz, *Sensations of Tone*, 1895, 109, for a comparative table of the pitches assigned the vowels by various investigators.

photometrically, and reached results which vary quite considerably from those of Helmholtz, whom Wolf follows in the pitches of vowels. The vowel clangs are found to be richer in distinguishable over-tones, than with Wolf, and variations in the absolute pitch of tonal elements are discovered. Recent work by Scripture[1] calls attention to the fact that the pitches are nearly always changing, and have values that vary with the interconnection of syllables, and the turns of expression given the words.[2] On the other hand, Bezold reaches results confirmatory of the earlier work of Helmholtz and Wolf as to the pitches of vowels and consonants. Doubtless, there will be restatements of these values, perhaps in terms of the variable pitches to which Scripture's results call attention, before final agreement is reached; yet for the purposes of tests one is safe in following Wolf and the later otologists who have developed and applied his results. In our present enquiry, the selection of speech tests, the following facts, established by Wolf, are important:

(1) Vowels carry farther than consonants.[3] Words with the vowels of light clang-tint,—a, e, i,—are more easily understood than those with the vowels of dark clang-tint,—o and u.[4] (German sounds).

(2) The consonants h, l, m, n, and w may be disregarded in speech tests, since they have no tones of themselves, but depend upon the vowel preceding or following.[5]

(3) The consonant sounds to which special regard must be paid in making tests are grouped by Wolf as follows. The range of each element in loud speech is also given:

I group:	High and strong:	Sch, 200 paces.
		S, 175 paces.
		Soft-G, or Ch-soft, 130 paces.
	High and weak:	F, 67 paces.
II group:	Medium pitch, the explosives.	
		K and T, 63 paces.
		B, 41 paces.
III group:	Deep sounds.	
	Lingual-R,	41 paces.
	[U (whispered, in diagnosis)	50 paces.]

[1] E. W. Scripture: *Studies of Yale Psych. Lab.*, 1899 and 1902; and *Philosoph. Studien*, xix, 599.

[2] *Zeits. f. Ohrenheilk.*, xxx, 1897, 114.

[3] Stated also by Helmholtz, *op. cit.*, 68.

[4] Wolf: *Sprache u. Ohr*, 96.

[5] Wolf: *op. cit.*, 15. Wolf's statement may be overstrong. In words, these letters of course take on pitch values, so that Scripture assigns them pitches as other sounds. Doubtless, too, they have some tonal elements of themselves.

These are the three characteristic groups of consonants in Wolf's system. Any adequate speech test must examine the perception for sounds in all three classes. Since the vowels are of farther range than the consonants, testing is based on the hearing of consonant sounds. Regard need be paid to the vowels only to include those of dark as well as of light clang-tint. There are other considerations in addition to the necessary inclusion of consonants and vowels of the classes just mentioned, which determine the selection of the actual words or other speech material to be used in the tests.

This material might include (*a*) the sounds of the letters pronounced separately, (*b*) these sounds joined into nonsense syllables, (*c*) sounds joined into words, or finally (*d*) words joined into sentences. Wolf found the simple unconnected sounds of the letters desirable in diagnostic tests: of these we shall speak later. In the testing of acuity, which we are considering here, sentences are objectionable, since the person tested may hear certain of the sounds or words, and then, aided by inflection and emphasis, only too easily guess at the whole expression. Words are open to the same objection, though in a very slight degree, since inflection, emphasis, and context are lacking. Nonsense syllables and the disconnected sounds of the letters avoid this difficulty, but they are open to the weightier objection that they do not directly test hearing for conversational speech,—our aim in this determination. All things considered, single disparate words recommend themselves as the material for the tests.

Granted that words should form our testing material, there is the further problem of the particular words to be chosen. In such a choice, regard must be paid to two conditions: (1) the structure of the words, *i. e.*, the vocal sounds included in the words and the interconnections in which the elementary sounds stand; and (2) the meaning of the words chosen. Under the first condition, provision must be made, as we have seen, for consonants of Wolf's three classes, and as far as possible all other elements of speech, including vowels of dark and light clang-tint. In addition, the following minor points[1] should be observed as far as possible: i. Words beginning with two consonants present difficulties in defective hearing and may well be included in the series of test words. ii. Perception for a consonant sound varies slightly, in some cases considerably, according as it appears as the initial, final, or emphatic syllable in a word. This point is of importance in diagnostic tests rather than in this general test; nevertheless, words with consonants in a variety of positions are desirable. iii. A regular

[1] Wolf: *Sprache u. Ohr.*, 95-96.

alternation of vowels and consonants in a word facilitates its perception.

The second condition limiting choice of test words, that of their meaning, is based on the fact that the ease of perception of a word varies directly with its familiarity. This is a very important factor. In ordinary conversation, we apperceive much more than we actually hear.[1] As already stated, this makes sentences undesirable for testing. With single words, even, it is unavoidable that to some indeterminate extent the test word will be apperceived when only part of its constituent sounds are heard. The determining factor in apperception of a disparate word from partial hearing is the degree of familiarity of the hearer with the word. An illustration will show where this fact leads us. The word "electricity," for example, other things equal, will be perceived accurately at a greater distance by a student of physics than the word "calico." The former word fits into his most common thoughts, and the least auditory suggestion of its sound is sufficient to evoke the verbal idea. The latter word, on the other hand, is an unusual one in his circle of thought, and a distinct perception of its auditory symbol is necessary to call the verbal idea to mind. To assume to compare the range of hearing for one word with that for another, as is necessary in testing a single individual, it is evident that we must have a series of test words with which the person tested is uniformly familiar, words whose sounds are equal in apperceptive value, *i. e.*, in calling up the verbal ideas corresponding to them. Let us follow our illustration farther. While a student of physics would understand the word "electricity" at a greater distance than the word "calico," a dry goods clerk would reverse the relative facility of apperception, and hear the word "calico" relatively farther than the word "electricity." As it is necessary to secure comparable results from various individuals, in order to determine the normal range for hearing and evaluate individual variations from it, test words must be uniformly familiar not only to the single individual, but, as nearly as possible, uniformly familiar to all individuals tested. This second factor, that of uniform apperceptive value, might be satisfied in either of two ways. (1) In a negative way, by taking vocal sounds entirely unfamiliar to the person tested; *e. g.*, the individual speech sounds given disparately, or nonsense syllables, or perhaps rare English or foreign words. The disadvantages of this are obvious: the tests would rapidly become unpleasant; unmeaning vocal combinations, whether in nonsense-syllables or unheard-of words,

[1] W. C. Bagley: *The Apperception of the Spoken Sentence, American Journal of Psychology*, xii, 80.

are intrinsically difficult to understand, for in ordinary life one pronounces a new word two or three times before feeling sure it is understood correctly; finally, such tests are too far removed from practical hearing to measure the latter adequately. (II) In a positive way, uniform apperceptive value may be sought by taking words in common use, known to every one, and, as far as may be, equally well known to every one. Number-words, the names of the numerals one to ninety-nine or higher, have been suggested as meeting this requirement.

Number-words present three distinct advantages which recommend them for testing purposes: (I) As a group of words, number-words are doubtless the most uniformly known, with any individual, of any series that could be devised. 52 is as familiar as 26; 13 as 8, etc. (II) Of all words, the number-words are doubtless the most uniformly familiar from person to person, despite differences of age, occupation, etc. (III) They are understood by children easily and with practical uniformity. They are the first body of words systematically taught, and are of daily use in school. Hence number-words are especially suitable for testing the hearing of school children, even in the lowest grades. Allowances must, of course, be made in actual tests for younger children who can count to ten or twenty only.

In view of these advantages, which ensure comparable results in the individuals, and between individuals as well, number-words commend themselves strongly as the material for tests. This statement, however, looks only to the second condition of choice, the apperceptive value of words. Do number-words satisfy as well the first criterion; do they include the various elementary sounds, both vowels and consonants, in sufficient number and variety of arrangement, to give adequate testing material? The series, nine to ninety-nine, is obviously limited. There are within it twenty-seven primary words which, either standing alone as "five" and "thirteen," or forming double words, as "twenty-seven" and "fifty-six," give the ninety-nine different number words in the series. It will be to the point to examine the primary number-words with regard to their consonantal and vowel sounds. Webster[1] gives the following pronunciation, to which in some cases are added German equivalents[2] for convenience of reference to Wolf's standards:

1.	wŭn		6.	sĭks	
2.	tōō	ōō = approx. Ger. ū	7.	sĕv'n	
3.	thrē	ē = Ger. i or ie	8.	āt	ā = Ger. ē
4.	fōr		9.	nīn	
5.	fīv	v = Ger. w; i = Ger. ei, ai	10.	tĕn	

[1] *International Dictionary*, 1903.
[2] C. H. Grandgent: *German and English Sounds*, Boston, 1892.

11.	ĕlĕv'n	20.	twĕntў ў = Ger. ĭ, also Eng. ĭ.
12.	twĕlv		
13.	thĕrtēn	30.	thĕrtў
14.	fōrtēn	40.	fôrtў
15.	fĭftēn	50.	fĭftў
16.	sĭkstēn	60.	sĭkstў
17.	sĕv'ntēn	70.	sĕv'ntў
18.	ātēn	80.	ātў
19.	nĭntēn	90.	nĭntў

An examination of these words shows the presence of the following speech elements, the accompanying numerals indicating the number of times each element appears:

Vowels

a	3 ā
e	8 ē, 7 ĕ, 1 ė̆, 2 ẽ
i	4 ī, 5 ĭ, 8 ў
o	2 ō, 1 ô
u	1 ŭ
oo	1 o͞o

Consonants

8 f, 3 k, 2 l, 20 n, 6 r, 9 s, 20 t, 6 v, 3 w, 3 th (as in *thin*).

Thus there are included in the number series ten different consonantal speech elements, each present from two or three times to twenty times. Webster enumerates twenty-five consonant elements as belonging to the English speech. Those not included in the number series are: ch, b, d, g, h, j, l, m, ng, p, sh, th (as in *thy*), y, z, zh. The fact that the number series includes less than half the consonantal elements of speech seems at first to condemn it, as inadequate to form the basis of speech tests. The answer again is not so easy; the adequacy cannot be decided off-hand by a mere enumeration of the sounds present. If these sounds include types of all consonant elements, they may very well test adequately the hearing for all. Comparing the consonants of the number series with the classification of consonant elements given in Webster,[1] one notes that all classes of consonants find representation by one or more sounds in the number series except the momentary surds (b, d, j, g). The final decision, however, whether the number series has sufficient representatives of all types of consonants, must be based, not on a classification like that of Webster's, made with regard to the pronunciation of speech, but on a classification made with regard to the hearing of speech. The consonants of the number words must be referred to Wolf's

[1] *Op. cit.*, lxvi.

three-fold classification, and the adequacy of the number series judged by it. The following Table shows under which of Wolf's classes each number-word falls; if the word contains consonant sounds of more than one class, it is included under each such class:

LOW.	MEDIUM.	HIGH.
2	2	
3		
4		4
		5
	6	6
		7
	8	
	10	
	12	
13	13	13
14	14	14
	15	15
	16	16
	17	17
	18	
	19	
	20	
30	30	
40	40	40
	50	50
	60	60
	70	70
	80	
	90	

Three number-words, one, nine and eleven, do not contain consonants falling within Wolf's classification. This does not, of course, exclude them from use in the tests. Considering the words falling within the three classes, the primary number-words are found to be distributed among the three classes as follows:

	LOW.	MEDIUM.	HIGH.
No. of words exclusively in each class	1	6	2
No. of words in each class and in another class as well	3	8	8
No. of words in all three classes	3	3	3
No. of different words containing sounds of each class	7	20	13

Our examination of the twenty-seven primary words shows that they are well distributed in the three classes of consonants. With these primary words, are to be included, in selecting test

words, the double number-words made from the primaries, *e. g.*, thirty-one, fifty-six. The primary words and the double words give in all 96 words, exclusive of one, nine, eleven, which are at our disposal. This certainly will furnish adequate material for series of testing words, with few repetitions save as the primary number-words reappear, differently compounded, in the various double-words.

At this point, an objection that has been made against the use of number-words in tests must be considered. Politzer[1] urges that, as the tests involve the use of a few primary words, the numbers will be judged with increasing ease as the test progresses, and the subject will come to guess the words, by hearing the vowels only. This objection Politzer considers insuperable. It is answered, however, by experimental results. Bezold used number-words in a test of 1,981 school children and in other extensive tests, and employs them constantly in otological practice. On the basis of the results of his school test, he says that there is no evidence of increased ease in the perception of the words as the tests progress, and whatever error there is from this source is practically uniform, since all the subjects know that they are being tested with number-words.[2] The experience of the author indicates that practice is a considerable factor, at first, but one that quickly reaches its maximum. Politzer is doubtless right in his objection, but he overestimates its importance. The error is known in direction, fairly constant in different individuals, and from its quick rise and small magnitude practicably negligible. It is certainly preferable to the variable errors of unknown magnitude involved in the apperception of the miscellaneous words which we must use in case we discard number-words. It may be added that number-words are in common and successful use by European otologists.[3] The objection of Politzer may, therefore, be overruled.

We return to the problem of selecting test words. The next step is to arrange lists of number-words, each list to contain words representing all three groups of consonants, and as far as may be with vowels of both dark and light clang-tint, and with consonants in the various positions, initial, medium, and final. The following lists seem to meet these requirements. The lists, further, may be expected to give comparable results, since they are equivalent as regards the elementary number-words included in each, and as regards the relative prominence of the three classes of consonants in each.[4]

---

[1] Politzer: *op. cit.*, 117.
[2] Bezold: *Das kindliche Gehörorgan*, 5.
[3] Bezold: *Functionelle Prüfung*, 206.
[4] Each of the ten lists contains uniformly the following elementary number-syllables: 1, "one;" 1, "two;" 1, "three;" 2, "four's;"

I	II	III	IV	V	VI	VII	VIII	IX	X
6	84	19	90	25	14	8	52	73	24
29	69	53	7	13	31	93	35	41	95
42	17	34	39	46	9	27	64	16	62
87	92	28	62	7	65	60	81	95	49
53	33	97	84	54	98	15	6	57	80
94	26	45	21	70	76	74	19	38	71
70	50	72	56	91	40	36	78	20	16
35	75	60	75	83	23	49	40	89	3
18	48	3	43	68	52	82	23	64	58
61	1	86	18	92	87	51	97	2	37

Speech tests are ordinarily given in whisper tones. In cases of hard-hearing, and sometimes in ordinary testing, conversational tones or even louder tones are employed. The merits of these two degrees of voice intensity are as follows.

Conversational speech (1) measures unequivocally the auditory functioning of practical life. Infrequent cases occur in otological practice in which whisper speech is heard farther than conversation (Burkhardt-Merian); and the whisper range is not a definite function of the range for conversation (Politzer). (2) Conversational speech must be used in determining cases of hard-hearing.

Whisper-speech (1) has approximately one-third the range of conversation and so permits testing in more restricted quarters. (2) In whisper-speech the vowels are reduced in intensity, while the consonants are little changed, so that the intensity of the vocal elements is more uniform than in conversation. This reduces the likelihood that the words will be guessed when only the vowel sounds are heard. On this score, whisper-speech has a distinct advantage, particularly when a limited number of test-words, as the number-series, is used. (3) Whisper-speech practically tests hearing for conversation. For these reasons, whisper tests are desirable and generally adequate. Limits of space, indeed, will usually preclude the use of conversational speech tests.

2, "five's;" 2, "six's;" 2, "seven's;" 2, "eight's;" 2, "nine's;" 1, "thir—;" 1, "—teen;" 1, "twent—;" 8, "—ty's." The relative prominence of the three classes of consonants in the ten lists, expressed as ratios of low consonants to medium to high (*i. e.*, l: m: h), is as follows:

I, 4:11:8; II, 5:12:8; III, 5:11:8; IV, 5:11:8; V, 5:12:9; VI, 5:12:8; VII, 5:12:8; VIII, 5:12:8; IX, 6:10:8; X, 5:11:9.

The writer compared these frequencies of the number series with that obtaining in the miscellaneous words of conversational speech and literary English. These latter showed the three classes of consonants present in the ratio 6:12:8, which is practically identical with the ratios given by the lists of number-words above.

Before considering actual methods of testing, the following important preliminary points of procedure may be stated.

1. Words are more easily perceived when the hearer is accustomed to the speaker's voice. This desirable familiarity will be secured by the preliminary explanations and statements made regarding the tests, which should be supplemented by a trial series of whispered test-words given according to the method of the actual tests. This serves the further purpose of making the subject feel at ease during the testing. This last is a very important point with children, and scarcely less so with many older people.

2. Two seconds or thereabouts before each test word, a warning signal should be given to insure maximal attention when the word is pronounced. Some uniform mechanical noise, as the sound of a pencil struck on wood, makes a satisfactory signal. The customary word "ready," pronounced in loud tones, may be used; but this necessitates an adjustment of the vocal organs for conversational speech just before each whispered test word, and is likely to interfere with uniform intensity in the latter. The signal should be repeated to indicate that the word has been spoken.

3. In testing the hearing for one ear, this ear is turned toward the examiner and the other ear closed. This can be most conveniently accomplished by pressing the tragus with the finger against the mouth of the meatus. Or, stoppers may be employed: Wolf used a rubber stopper; Politzer, the moistened finger in the meatus, or an olive-shaped obturator; and Titchener recommends the cap-shaped rubber eraser (found on many lead pencils), filled with laboratory wax before using. Small corks softened with vaseline are also quite satisfactory. When loud words are spoken directly before a diseased ear in hard-hearing tests, one must be sure, when words are heard, that the ear on the other side is not functioning despite its closure. Dennert suggested the check test of stopping both ears and repeating the words before the diseased ear. If words are still heard, the results of the preceding tests are unreliable, as the perception in both tests was doubtless mediated by the other ear; if words are not heard with both ears closed, the former test of the diseased ear was reliable.

4. The subject must not see the lips of the examiner in the tests, as slightly deaf persons often unconsciously "read the lips." In tests of one ear, the subject sits sidewise to the examiner and his position in part guards against this error. It is safer to have the eyes closed or shielded; the mouth should be closed through the tests, as the intensity of sound varies according as the mouth is open or closed. Such precautions

should be observed constantly. Any intermission interferes with regular attention.

5. With children, special care is necessary. As suggested above, conversation about the test and friendly treatment will assist in putting them at ease. If inattention and restlessness intervene, the tests should be stopped and resumed later.

6. As far as possible, variations in the voice of the examiner must be avoided. An approximately even intensity may be secured by always pronouncing test words with the residue of air in the lungs after the normal expiration (Bezold). If the test words are given successively, there may be interpolated after each word a definite number of breathings, three or four, the following word being pronounced after the last expiration in the group of breathings. Regard must also be paid to securing uniform tempo and distinct articulation. It is impossible to rule out differences in clang-tint between different voices. The variations in the individual voice may be largely obviated by the precautions suggested, and with practice it attains a closer approximation to uniformity.

7. It is generally desirable to examine each ear separately and then to test binaural hearing. The binaural range (*a*) is often unrelated to those of the ears functioning separately; and (*b*) it is precisely this range which is of importance in daily life. In the binaural test, the subject faces the examiner and closes his eyes.

8. The testing room for whisper speech should have a range of 30 m., if possible. Tests may be made, however, with an extreme range of only 18-20 meters, and with the second method suggested below even shorter ranges are satisfactory. The normal whisper range given by various investigators, varies from 17-41 m.[1] These different values are doubtless ascribable to the use of whispers of different intensities, and to the varying acoustic conditions in the various tests. The acoustical error might be obviated by making tests in the open air; but such testing is not commonly practicable. By performing tests in rooms of average oblong shape, and noting dimensions and other matters of acoustical importance, as disturbing noises, investigators will add to the comparability of their results. Bezold found that bright sunlight streaming from a window across his testing room changed its acoustic properties. The room should be kept free from noise within and without. Uniformity

[1] The following ranges have been given as norms by various investigators: Hartmann and Siebenmann, 25-26 m. in a quiet room and 20 m. in a room disturbed by the noise of the day; O. Wolf, 20 m.; Chimani, with soldiers, 21 m. (Politzer, *op. cit.*, 116.) F. Bezold, in his tests of school children, 17-20 m; Matte and Schultes, testing soldiers with a medium strong whisper, 35-40 m.; *Zeits f. Phys. u. Psych.*, xvi, 308.

should be observed as to position of large pieces of furniture, opened and closed windows and doors, and other conditions affecting sound reflection. Tests for conversational speech require a range three times that for whisper-speech (Bezold), and so a minimum space of perhaps 50 m. A range-line should be chalked on the floor in the direction of the greater dimension of the room, passing through the center and scaled to meter and half-meter divisions. It may be supplemented, if necessary, by a meter stick for short ranges in hard-hearing. The examiner, in working up and down the range, should always face the subject squarely as he pronounces test words. The subject will have a constant position at the end of the range-line, where he should be seated comfortably. If the testing is by groups, the individuals may be seated along an arc of 15 m. radius cutting the range at right angles.

Actual Performance of Tests. If a group of subjects are to be examined, rough preliminary tests may be made by pronouncing number-words at various ranges, on the results of which the group may be divided into two or more sub-groups, each containing persons of about the same grade of acuity of hearing. This will facilitate the actual testing. The precautions already suggested are to be observed: familiarity with the speaker's voice and with whisper-speech, closure of the ear not under test, closing of subject's eyes and mouth, attention secured by the warning signal just before the test-word, uniformity in the examiner's voice, and constant acoustical conditions in the testing room. With the individual or group seated at the end of the range-line, the examiner pronounces successively in whisper-speech a series of ten number-words, at each of the ranges selected, following this procedure: The warning signal, or "rap," is given, followed at the conclusion of the next breath by the first test word, which is pronounced with the residual air in the lungs. It is followed by a second warning "rap," signifying to the subject that the word has been pronounced and that he may now open his eyes and write down the numerical symbol of the word heard, *e. g.*, 22 or 9. If he has heard nothing, he indicates this fact by a dash (——). Meantime, the examiner interpolates three complete breaths and then gives the warning signal preparatory to the second test-word. At the signal, the subject closes his eyes and sits at attention until the examiner at the end of the next breath gives the second test-word, indicating its conclusion by another signal. Thus there follow successively: three complete breathings, the preparatory signal, the test-word at the conclusion of the next breath, the concluding signal, and so on through the series of ten words. This is the general procedure in the tests. A further statement must be made of two methods, either of

which may be followed in the tests : (1) the common method of measuring auditory acuity by the extreme range of accurate hearing ; (2) the author's method of measuring acuity in terms of degree of accuracy of hearing at some constant range or ranges.

Method of Extreme Ranges. In this method the examiner makes rough preliminary tests with a few whispered number-words to determine the subject's approximate extreme range. An accurate determination of the audible range is then made by beginning at a point on the range-line 3-4 meters beyond this approximate range and moving in by meter or half-meter intervals, giving a series of ten test-words at each point. The point at which the subject hears correctly some arbitrary part, say 80%, of the list of ten words, is regarded as the individual's extreme range. The average of the ranges of all the individuals tested under the same conditions is the normal range. The individual's auditory acuity is expressed as a fraction whose numerator is his own range and whose denominator is the average or normal range.

The propriety of this method depends on the assumption that the intensity of the sounds of speech decreases with approximate regularity as the distance from the speaker increases. The author's experimental use of speech tests has led him to doubt the truth of this assumption. As set forth in the appendix, his results indicate that in a closed room there is no regular scale of decreasing intensity as the range increases. There is a general decrease in intensity as the indoor range increases, but the reflection of sound-waves back and forth in the room makes almost meaningless such distinctions as that A hears 80% correctly at thirty feet, B at forty feet, and C at fifty feet. It may happen, as it did in the author's work, that it is paradoxically easier to hear words in the particular room at fifty feet than at forty feet. The existence of this source of error throws no little doubt, in the author's opinion, upon the work which has heretofore been done in speech tests. Further experimental testing is needed before final conclusions can be stated regarding the extent to which the above method is trustworthy. Meantime, the following second method of procedure is advanced with the design of avoiding the error due to reflection of sound-waves, which seems to be inherent in the above method.

Method of Degree of Accuracy. In this method, one, two, or more, constant ranges are selected according to the testing space at the disposal of the examiner. If a single range is used, it should be on the border line of difficult hearing; if three ranges are chosen, one may be at easy hearing distance, one at a moderately difficult distance, and one at a quite difficult distance.

All ten testing lists, *i. e.*, the whole hundred testing words, are then spoken at the single range; or if more than one range is used, the words are divided about equally between the two or three ranges. Each individual is examined with the 100 words, and his results are computed as to the percentage of accurate audition. The norm is taken as the average of all the individual percentages; and each individual's audition is then expressed as a fraction of this norm. If time is valuable, the test may be shortened to fifty words or even less. If space is restricted, the range may be shortened by the use of screens to cut off the sound-waves and so reduce the intensity of the sounds. This may be easily accomplished by the examiner taking his position in an adjoining room with the connecting door open. It is necessary with this method, and especially so with artificially shortened ranges, that the examiner and the subject keep certain constant positions, even with regard to the direction of the head, throughout the entire series of tests.

The advantage of this second method lies in its freedom from the errors due to the reflection of sound-waves. With the examiner and subject in certain definite positions at a certain constant range, or ranges, the acoustical conditions remain uniform throughout the tests; and if A hears 80% accurately and B 60% accurately, it may be concluded with apparent certainty that A's hearing is more acute than B's. If the average accuracy of all persons tested, moreover, is 70%, A may be regarded as of super-normal acuity, and B as of sub-normal acuity. The theoretical superiority of this method is supplemented, too, by certain experimental results which, while not furnishing final proof, do in certain instances point toward the correctness of this method. These are given in the appendix.

Summary of Speech Tests. This discussion of speech tests has urged their importance as the only unequivocal measure of the ability to hear speech. An analysis of speech into its elements (following Wolf), and an examination of the apperceptive conditions of audition, furnished the criteria by which the material for speech-tests was selected. Judged by these criteria, number-words were found to recommend themselves for use in tests. Ten comparable and equivalent lists of ten number-words each were selected. The general conditions of conducting speech tests were given, followed by a statement of two methods of testing: the first, the common method of measuring the individual's acuity in terms of his extreme range of audition; the second, a proposed method of pronouncing all the test-words at one or more common constant ranges, and determining the individual's acuity in terms of the percentage of accuracy of audition.

B. Tests of Acuity of Hearing with Tones and Noises.

While speech tests alone measure unequivocally the ability to hear speech and, therefore, are to be employed, if possible, to determine acuity of hearing, limitations of space and other difficulties of using speech tests will often require the employment of tests with mechanically produced sounds. Reserving for a later section those methods which have their particular significance in the diagnosis of aural dieases, we will consider here, I, the four chief methods of testing acuity of hearing with mechanical sounds,—(*a*) with the watch, (*b*) acoumeter, (*c*) audiometer and other electrical devices, and (*d*) with forks; II, the tests of the lower and upper limits of tonal audibility.

I. Tests of Acuity of Hearing. It is a familiar fact that individuals vary in sharpness of hearing. Hard-hearing is an almost invariable accompaniment of old age,[1] and many individuals, young and old, have sub-normal keenness of hearing through innate conditions, or as a result of disease. The disturbance ordinarily manifests itself in difficulty in conversation, though, if the deficiency is slight, the person affected and his friends may alike be ignorant of its existence. The detection of such cases is particularly important in the public schools, where pupils who are regarded as stupid are often only hard of hearing, and could be properly provided for by advantageous seating and other special attention, were the real nature of their trouble understood. Tests of auditory acuity should also be applied to railroad men, telegraph and telephone operators, the police, soldiers, and all individuals in public or private service of a nature that demands accurate audition. In part, speech tests can be used for the purpose; but tests more convenient of application are often demanded, even if their results are not so unequivocal. Further, tests of finer adjustment than speech tests are needed to detect slight reductions in acuity of hearing or slight improvement in cases under treatment.[2] The tests with mechanical tones and noises meet these requirements. We shall consider several such tests in order.

A. Watch Test. This has long been the most common test of hearing and is still the one most frequently used in medical practice. The objections to the watch test have already been partly set forth in treating speech tests. Its results are not a

[1] Bezold's results give a comparison of hearing in school-children, and in persons over fifty years of age, as follows: 46.5% of the school-children examined had a range for whisper-speech of 16 m. or more; only 15.57% of persons fifty to sixty years old had a range of 8 m. or more, and only 6.7% of those from sixty to seventy had this range. Not only is there a decrease in acuity in old age, but the decrease becomes relatively more rapid as it proceeds. *Funkt. Prüf.*, 152-154.

[2] Politzer: *op. cit.*, 110.

reliable indication of the capacity to perceive speech; its sounds are a series of equidistant tones (Wolf) and so give rise to a perception of rhythm, not of simple sound; its ticking is so familiar that illusions of hearing are very frequent, *i. e.*, the subject imagines he hears the watch and reports it as heard even when it is beyond range; watches vary in the quality and intensity of their sounds, and hence results are not directly comparable.[1] On the other hand, the watch is always at hand and convenient to employ, and the procedure can be so arranged as to resolve, in part, the objections just stated. Nevertheless, the objections stand in part, and the writer recommends strongly that the watch be used only in cases where more accurate and standardized instruments cannot be secured.[2]

The procedure for watch tests consists essentially in determining the last point at which the ticking can be heard as the watch is carried out from the ear; then, after passing beyond the audible range, the watch is brought in toward the ear until again heard; the average of the two distances at which it is just heard is the range for the ear tested. The details of procedure may be stated as follows:

1. A range should be chalked on the floor 9-10 meters long, marked with half-meter sub-divisions from the one-meter point on. The meter stick itself should be used for distances less than a meter, being held for this purpose at right angles to the ear, and the watch moved along it. The range should extend along the middle of the room, and, as in the speech tests, the subject should be seated.
2. The subject's eyes should be kept closed, or better, covered with a blind. If he carries a watch, it must be laid aside. Quiet must be observed within and about the room.
3. The watch should be carried out along a line at the level of the ear and a little forward of a line passing through the two ears. It should always be held in the same relative position, preferably with its face toward the ear tested.
4. Tests should be made at the half-meter points along the range till a point is reached at which the ticking is no longer heard. The procedure is then reversed, and the watch brought in by stages, until again heard. The average of the two values is the range for the ear tested.

---

[1] F. Bezold states fully the objections to the watch in *Archives of Otology*, xv, 1886, 79; and on p. 9 of *Das kindl. Gehörorgan* gives comparative results showing the discrepancies between watch, acoumeter and speech tests. The acoumeter and voice agree better than the watch and voice. *Cf.* Tafel I, opp. p. 94, *op. cit.*

[2] Cattell and Farrand used the stop-watch (*Psych. Rev.*, iii, 635). It is also recommended by the Committee on Tests (*ibid.*, iv, 134). An ordinary watch, however, will serve as well if a stop-watch is not at hand. Stop-watches have been suggested by some with the method of stopping and starting the watch and requiring the subject to tell how many times he hears the watch going. The click made in starting and stopping most stop-watches is so loud, however, that the method is useless. This method can be used with an ordinary watch by covering and uncovering it in the hands.

5. Check tests should be constantly interpolated, especially with children, by covering the watch with the two hands and then asking if it is heard. This will disclose illusory judgments.

6. Each ear is to be tested separately, the other being meantime securely closed.

7. A normal range must be determined for the particular watch used. This can be taken as the average of a large number of individual determinations, and will vary with different watches. The hearing acuity of an auditory organ is then expressed as a fraction, whose numerator is its range, and whose denominator is the normal range; *e. g.*, if the average range is 2 m. and A's left ear has a range of 1.5 m., its acuity is ¾ normal. Expressed in such a way, the results of different investigators using different watches can be compared.

B. Acoumeter Tests. Dissatisfaction with the watch as a means of testing led Politzer to invent a small instrument, the acoumeter, which could be used as a substitute.[1] It consists of a small steel cylinder fastened at right angles into a hard-rubber upright, with a percussion hammer which can be raised a constant distance and allowed to fall upon the cylinder. This gives a tone (c^2), uniform in pitch and intensity. The instrument as described is arranged for air-conduction. It is also provided with a metal disc which can be set directly against the temporal bone, mastoid process, or other parts of the skull, if it is desired to test hearing with bone-conduction. The general procedure with the acoumeter, and the conditions to be observed, are precisely the same as in the watch test, and in making tests, reference should be had to the statements given in the preceding sections. In addition, the following points should be observed.

A large room will be needed for testing, for the instrument has a normal range of 15 m., which extends for very keen ears to 20 m. or more. The acoumeter must always be held perpendicular in giving the stimulus. The "ready" signal is given before each test, and fatigue avoided by allowing a short interval between successive tests, and limiting the length of a series. The actual procedure, in the receding series, consists of giving a varying number of acoumeter tones (1, 2, 3 or 4) at the successive points on the range, 10 m., 10.5 m., etc. The subject each time reports the number of tones he hears. The receding range is that at which he last reports correctly the number of tones given. The approximating series is made in a similar way, by starting at a point at which the sounds are inaudible. The range at which there is the first accurate report of the number of tones given, is the approximating range. The two are averaged for the range of the ear under test. The normal range is taken as the average of the various individual ranges, as with the watch, and the acuity of an individual ear is expressed as a fraction whose numerator is the range for the particular ear, and whose denominator is the average range; thus the results of various investigators are made comparable.[2]

---

[1] Politzer: *op. cit.*, 107-8.

[2] Politzer and others give the normal acoumeter range as 15 m. Bezold's results (*Das kindl. Gehörorgan*, 7) indicate a normal range of 16 or 17 m. This emphasizes the necessity of computing the normal

The acoumeter gives a more reliable measure of relative acuity for hearing speech, than does the watch.[1] It is, therefore, to be recommended instead of the watch. Its range is approximately the same as that for whisper-speech, and accordingly it has no advantage over speech tests in the space required for its employment. It should be borne in mind that the ranges of hearing with the watch, acoumeter, audiometer, etc., have no psychological significance, and are not to be confused with psychophysical limens.

C. The Audiometer and other Electrical Devices. Many attempts have been made to adapt electrical devices as testing instruments. The general plan has been to use a telephone click which will vary in intensity according to the amount of electric energy employed in producing it. The amount of current so used has been varied by a sliding induction coil, or by adjustable resistance, and admits, of course, of ready measurement. This, in turn, evaluates the intensity of the sound produced. Hughes and Boudet, Hartmann, Blyth, Koerting, Urbantschitsch, Preyer and others have devised and used various instruments of this sort.[2] Politzer makes the general objection that such instruments do not provide accurate measurement of sound intensity. This difficulty is apparently resolved in an audiometer recently devised by Seashore.[3]

This instrument gives a double telephone click, whose intensity can be varied through forty stages which are approximately equal in consciousness. The instrument includes a dry battery, primary coil, variable resistance coil, variable secondary coil, and a galvanometer, with necessary switches and wires leading to a telephone receiver. The primary current, by means of the variable resistance, can be adjusted to remain at a certain constant amount. The induced current may be varied in intensity through the forty steps, which furnish a measure of the intensity of the click given in the receiver on making and breaking this current. This instrument can be used with limited floor space, and outside noises give less disturbance than with most testing methods. An ascending and a descending series of tests are made, and the results averaged. Check tests should be inserted, especially when near the limen of hearing, by giving the "ready" signal and then omitting the stimulus. The two sounds produced by the Seashore audiometer are not, at least not always, alike in quality. This objectionable feature could doubtless be eliminated by a modification of the apparatus. The results are in terms of the arbitrary scale of forty steps in intensity; but they can be transposed into a more comparable

(average) range for the particular instrument used, and the particular conditions of each test. Politzer calls attention (109) to variations in individual results on different days, at different hours of the day, and under different physical and mental conditions (fatigue, anxiety, inattention, etc.). These facts should be remembered in employing the acoumeter, and, indeed, in all testing.

[1] Politzer: *op. cit.*, 109; Bezold: *op. cit.*, Tafel I, opp. p. 94.

[2] Gruber: *Diseases of the Ear*, 131; Politzer: *op. cit.*, 110.

[3] For a full description, see *Univ. of Iowa Studies in Psych.*, II, 158. This instrument is sold by the Chicago Laboratory Supply Co.

and intelligible form by writing them as fractions of the normal (average) result given by all the individuals tested.[1] The audiometer is also arranged to give a tonal stimulus. The tone may be varied and measured, Seashore says, in the same way as the noise stimulus. The author has not had an opportunity to test the audiometer in practical work.

D. Acuity Tests with Forks. The tuning fork is not adapted for acuity tests based on the distance at which a tone can be heard, since the tone given by a fork sounds first at a high intensity and then gradually "dies off;" when first struck, accordingly, it can be heard at a greater distance than a moment later when its tone has partially rung off. Van Conta[2] proposed a temporal method for forks which compares the ringing-off time for the ear under test with that for another ear used as a standard. It is based on the fact that the greater the acuity of hearing, the longer the tone will be heard.

A tuning fork (c^2) is struck and brought before the ear to be tested and when no longer heard, is transferred to the examiner's ear and the time it continues to sound noted. If necessary the procedure is reversed: the fork is placed first before the examiner's ear, and then before the ear under test. The obvious error in the variable time consumed in transferring the fork is avoided by Urbantschitsch, who uses a T-shaped tube with one branch in the subject's and one in the examiner's ear. The fork is held before the free end of the tube. By the use of two stop-watches each person can indicate accurately the length of time the tone is heard. There remains the error due to varying initial intensities of the tone, which can be reduced, but not eliminated, by care in striking the fork. This test obviously has a limited application.

II. Tests of Upper and Lower Limits of Hearing. Tonal sensations are evoked by the impact of periodic sound-waves upon the organ of hearing. Air waves of less than a certain minimal frequency do not produce a tonal sensation; similarly, air waves of more than a certain maximal frequency do not give rise to tones. We shall consider tests measuring the individual's upper and lower limits of tonal sensation. Such tests cannot pretend to the nicety of method or accuracy of results aimed at in psychological determinations. It is possible, however, to make rapid approximate measurements which will reveal individual variation in these capacities.

A. The Lower Limit of Tonal Sensation. Two distinct paths have been followed in determinations of the lower limen: it has been approached by trained investigators with skilled observers, and all the niceties of scientific method; so we have

[1] *Child Study Report, No. 2, 1899-1900, Chicago Public Schools,* gives results from testing 5,706 pupils with the audiometer. *Cf.* pp. 60 ff. See also *Some Results of Hearing-Tests of Chicago School Children,* D. P. MacMillan, *Medicine,* April, 1902.

[2] *Archiv f. Ohrenheilk.,* I, 107.

the work of Helmholtz, Wundt, Preyer, Appunn, and others who have secured results varying from 8 to 28 vibrations as the lowest audible tone.[1] Measurement of the lower limen has also been used in otological practice and in tests by Bezold, Zwaardemaker, and others, upon patients and other persons who are without training in scientific observation, and who present all extremes of general intelligence. Our proposed test plainly falls in with the latter rather than the former class. Of the instruments used in lower liminal determinations, the lamella and the wire forks are the most practicable for wide-spread use in tests. Testing methods will be stated for these two instruments.

(*a*) The lamella consists of a long, thin, steel blade, 420 mm. by 12 mm. by 1 mm., with a metal disc at its upper end. It is clamped in a wooden vise which in turn is fastened to a table. The lamella has a scale marked along one face showing the point to which it should be inserted in the vise to give 4, 6, 8, 10, ... 24 vibrations per second. When clamped in position at the desired point on its scale, it is put into vibration by springing it from a position of equilibrium with the finger. Thus sound waves of from 4 to 24 vibrations a second can be produced. Overtones are eliminated by adjusting a band of cloth one-third of the distance from the upper end of the blade.

There are two sources of error. (I) An error of observation. The subject is likely to mistake the discrete puffs of air at the lower rates for tones (Bezold). Practice should be given with a few preliminary tones at the higher rates which are usually clearly distinguished as tones, and then with the lowest rate which will give discrete puffs of air; and the subject's attention called to their qualitative difference. (II) An objective error. The length of the lamella increases as the tone deepens, and hence the amplitude of vibration and the intensity of the tone increases (Zwaardemaker). This cannot have much weight, as the difficulty of perception increases as the tone deepens, and the facility due to increased intensity may be regarded as an offset to the greater difficulty of perception. The procedure with the lamella involves the usual ascending and descending determinations and an averaging of results.

(*b*) The Appunn wire forks are made of wire bent in the usual fork form, and provided with discs on the ends of the prongs. The series includes forks differing by two vibrations from eight to twenty-four vibrations, and by eight vibrations from twenty-four to fifty-six vibrations. These forks present the following advantages: approximate freedom from overtones;

[1] Helmholtz used organ pipes and a vibrating monochord: *Sensations of Tone*, 1895, 175-6. Preyer used reeds and a heavily weighted fork: *Die Grenzen d. Tonwahrnehmung*, Jena, 1876. Wundt used the heavy fork, the Appunn lamella or vibrating blade, and difference tones: *Grundzüge d. phys. Psych.*, 4th ed. I, 450; other investigators have used Appunn's wire forks.

tones loud enough to be audible; the initial amplitude is constant for each fork; they are easily handled, and present no serious difficulties for untrained subjects. The series should be applied in both orders, and an average struck between the results.

With the lamella, Zwaardemaker found average values for the lower limen for different ages at from 10 vibrations to 13 vibrations.[1] Bezold, working with loaded forks (Edelmann's), uses a lower limen of 16 vibrations, though for many persons, he says, this doubtless lies above the actual limen. The general results of tests will probably agree more closely with Zwaardemaker's figures.

B. The Upper Limit of Tonal Sensation. The most practical instrument for upper limit tests is the Galton graduated whistle. This was invented by Francis Galton and has been recently improved by Edelmann.[2] It consists of a covered cylindrical organ-pipe of small dimensions with a small air-bulb as bellows. The pitch of the whistle is varied by a micrometer screw which moves a piston up and down the cylinder, thus altering the length of the pipe. Readings on the micrometer scale correspond to the pitches to which the whistle can be adjusted. As modified by Edelmann, it has a second micrometer screw which adjusts the lips of the whistle to give an optimal tone at the various pitches. A schedule accompanies each instrument showing the proper setting of the two adjustments for the pitches within its compass. One form covers the tonal scale from a^4 of 3,480 vibrations to f^9 of 44,193 vibrations and above that to 49,000 vibrations. A scale of

[1] A. Zwaardemaker: *Zeitsch f. Psych. u. Phys.*, vii, 1894, 10. Zwaardemaker in this article attempts to establish a correlation between increasing age and a shortening of the tonal scale. The average lower limit for ten to twenty years is 10.10 vibrations; for thirty to forty years, 10.85 vibrations; over sixty years, 12.95 vibrations. Tonal audition extends over eleven octaves in youth, he concludes, and only ten octaves in old age. He reached these results with the lamella and the Galton whistle. Bezold (*Funkt. Prüfung*, 144), using the weighted forks and the whistle, finds that the shrinkage is much less, and explains it as due to injuries and disease, rather than advancing age.

[2] Determinations of the upper limit have also been made in the past with small forks, made by Appunn, Koenig, and Edelmann, and with Koenig's cylinders. Appunn's forks consist of one series of eleven forks, varying by steps of 5,000 vibrations, from 5,000 to 45,000, and another series of thirty-three forks from c^4 (2,408 vib.) to g^8 (49,152 vib.). Discredit has been thrown on determinations with Appunn's forks through recent investigations showing large errors in their vibration markings. (For statement of discussion see *Zeitsch. f. Psych. u. Phys.*, xxi, 141; xxii, 229; xxiv, 171 and 367.) Koenig's cylinders are of steel, 20 mm. in diameter, and of varying length. One set of ten cylinders covers the range from 4,096 to 32,768 vibrations. There is also another series of 22 cylinders covering the same interval.

intervals by steps of 1,000 vibrations is also usually given, from 10,000 vibrations to the upper limit. Zwaardemaker notes two advantages of the whistle: it is approximately constant in intensity,[1] and its tone is prolonged but an instant, and so does not fatigue the ear. Further, it may be said that the whistle is simple to operate, and gives a tone which when present is not difficult to distinguish from the accompanying noise. In testing with the instrument, an ascending and descending determination should be made and averaged. Checks should be made on judgments at the limen, by interpolating occasional pure noises between the tones. Results may be expected to vary considerably; it should be remembered that large differences in vibration rates at 30,000 to 50,000 vibrations mean only a small tonal difference; *e. g.*, the interval just stated, which comprises 20,000 vibrations, is only a major 6th.[2] Perception at the upper limen is very fatiguing, and series must accordingly be limited.

APPENDIX.

An Experimental Examination of Speech-Methods.

We have given in the text what may be called the 'traditional' method of testing audition by means of speech. This method, it will be remembered (see p. 40), is the method of Extreme Range of Hearing. A normal maximal distance is selected and, on the basis of it, individual capacities are computed by comparing the individual ranges with the normal range. Thus, if the normal range is twenty meters and the individual's limit falls at ten meters, the individual's capacity is represented by some fraction: *e. g.*, ½.

Now it is evident that the correlation of auditory capacity with distance will be valid only on condition that distance and audibility stand to each other in some simple functional rela-

[1] E. W. Scripture: *Studies of Yale Psych. Lab.*, II, secured constant pressure by mechanical means. See, however, Myers' results in the next foot-note.

[2] Edelmann gives 50,000 vibrations, or g^6, as the upper limit; Schwendt, e^5-f^5 (41,000-44,000 vibs.); Zwaardemaker, with persons of all ages, from c^6 to f^7 (22,000 vibs.). For Edelmann, see *Zeitsch. f. Ohrenheilk.*, xxxvi, 1900, 330; Schwendt, *Pflüger's Archiv*, lxxv, 346, lxxvi, 189, and *Archiv f. Ohrenheilk.*, xlix, 1; Zwaardemaker, *Zeitsch. f. Psych. u. Phys.*, vii, 10. Very recent determinations by C. S. Myers (*Journal of Physiology*, London, xxviii, 1902, 417-425) throw doubt on all high results with the Galton whistle. Myers concludes that the pitch of the whistle varies with the absolute air pressure and with changes in air pressure; and that the liminal value of 50,000 vibs. is unreliable. He puts the limen at 20,000 to 25,000 vibrations for young adults. Tests may be expected to give results like Zwaardemaker's and Myers', *i. e.*, about 20,000 vibs.

tones loud enough to be audible; the initial amplitude is constant for each fork; they are easily handled, and present no serious difficulties for untrained subjects. The series should be applied in both orders, and an average struck between the results.

With the lamella, Zwaardemaker found average values for the lower limen for different ages at from 10 vibrations to 13 vibrations.[1] Bezold, working with loaded forks (Edelmann's), uses a lower limen of 16 vibrations, though for many persons, he says, this doubtless lies above the actual limen. The general results of tests will probably agree more closely with Zwaardemaker's figures.

B. The Upper Limit of Tonal Sensation. The most practical instrument for upper limit tests is the Galton graduated whistle. This was invented by Francis Galton and has been recently improved by Edelmann.[2] It consists of a covered cylindrical organ-pipe of small dimensions with a small air-bulb as bellows. The pitch of the whistle is varied by a micrometer screw which moves a piston up and down the cylinder, thus altering the length of the pipe. Readings on the micrometer scale correspond to the pitches to which the whistle can be adjusted. As modified by Edelmann, it has a second micrometer screw which adjusts the lips of the whistle to give an optimal tone at the various pitches. A schedule accompanies each instrument showing the proper setting of the two adjustments for the pitches within its compass. One form covers the tonal scale from a^4 of 3,480 vibrations to f^8 of 44,193 vibrations and above that to 49,000 vibrations. A scale of

[1] A. Zwaardemaker: *Zeitsch f. Psych. u. Phys.*, vii, 1894, 10. Zwaardemaker in this article attempts to establish a correlation between increasing age and a shortening of the tonal scale. The average lower limit for ten to twenty years is 10.10 vibrations; for thirty to forty years, 10.85 vibrations; over sixty years, 12.95 vibrations. Tonal audition extends over eleven octaves in youth, he concludes, and only ten octaves in old age. He reached these results with the lamella and the Galton whistle. Bezold (*Funkt. Prüfung*, 144), using the weighted forks and the whistle, finds that the shrinkage is much less, and explains it as due to injuries and disease, rather than advancing age.

[2] Determinations of the upper limit have also been made in the past with small forks, made by Appunn, Koenig, and Edelmann, and with Koenig's cylinders. Appunn's forks consist of one series of eleven forks, varying by steps of 5,000 vibrations, from 5,000 to 45,000, and another series of thirty-three forks from c^4 (2,408 vib.) to g^8 (49,152 vib.). Discredit has been thrown on determinations with Appunn's forks through recent investigations showing large errors in their vibration markings. (For statement of discussion see *Zeitsch. f. Psych. u. Phys.*, xxi, 141; xxii, 229; xxiv, 171 and 367.) Koenig's cylinders are of steel, 20 mm. in diameter, and of varying length. One set of ten cylinders covers the range from 4,096 to 32,768 vibrations. There is also another series of 22 cylinders covering the same interval.

intervals by steps of 1,000 vibrations is also usually given, from 10,000 vibrations to the upper limit. Zwaardemaker notes two advantages of the whistle: it is approximately constant in intensity,[1] and its tone is prolonged but an instant, and so does not fatigue the ear. Further, it may be said that the whistle is simple to operate, and gives a tone which when present is not difficult to distinguish from the accompanying noise. In testing with the instrument, an ascending and descending determination should be made and averaged. Checks should be made on judgments at the limen, by interpolating occasional pure noises between the tones. Results may be expected to vary considerably; it should be remembered that large differences in vibration rates at 30,000 to 50,000 vibrations mean only a small tonal difference; *e. g.*, the interval just stated, which comprises 20,000 vibrations, is only a major 6th.[2] Perception at the upper limen is very fatiguing, and series must accordingly be limited.

APPENDIX.

An Experimental Examination of Speech-Methods.

We have given in the text what may be called the 'traditional' method of testing audition by means of speech. This method, it will be remembered (see p. 40), is the method of Extreme Range of Hearing. A normal maximal distance is selected and, on the basis of it, individual capacities are computed by comparing the individual ranges with the normal range. Thus, if the normal range is twenty meters and the individual's limit falls at ten meters, the individual's capacity is represented by some fraction: *e. g.*, ½.

Now it is evident that the correlation of auditory capacity with distance will be valid only on condition that distance and audibility stand to each other in some simple functional rela-

[1] E. W. Scripture: *Studies of Yale Psych. Lab.*, II, secured constant pressure by mechanical means. See, however, Myers' results in the next foot-note.

[2] Edelmann gives 50,000 vibrations, or g^8, as the upper limit; Schwendt, e^8-f^8 (41,000-44,000 vibs.); Zwaardemaker, with persons of all ages, from c^6 to f^7 (22,000 vibs.). For Edelmann, see *Zeitsch. f. Ohrenheilk.*, xxxvi, 1900, 330; Schwendt, *Pflüger's Archiv*, lxxv, 346, lxxvi, 189, and *Archiv f. Ohrenheilk.*, xlix, 1; Zwaardemaker, *Zeitsch. f. Psych. u. Phys.*, vii, 10. Very recent determinations by C. S. Myers (*Journal of Physiology*, London, xxviii, 1902, 417-425) throw doubt on all high results with the Galton whistle. Myers concludes that the pitch of the whistle varies with the absolute air pressure and with changes in air pressure; and that the liminal value of 50,000 vibs. is unreliable. He puts the limen at 20,000 to 25,000 vibrations for young adults. Tests may be expected to give results like Zwaardemaker's and Myers', *i. e.*, about 20,000 vibs.

tion, and further, on condition that the sounds used as test-words undergo, with change of distance, merely a quantitative, not a qualitative, alteration. These conditions have been assumed, without sufficient ground, by the traditional method.

The obvious way to test these assumptions is by appeal to experiment. Two groups of experiments were, accordingly, arranged and carried through in a manner presently to be described. In the first group, a large number of ranges was employed as a preliminary step to the selection of a final, maximal range, as provided for by the method. These experiments brought to light the fact that the two essential conditions just named are not to be found in the method as it stands.

The second group had, then, for its end the discovery of a more adequate mode of procedure.

I. *Preliminary Tests for the Determination of Extreme Range.* The experiments in Group I were conducted in the Armory Hall of Cornell University, a rectangular room 40 by 120 feet. The range began at a point twenty feet from one end of the room, and was laid off along the median line for a distance of eighty feet. The person giving the test-words stood at the zero point of the range, while the observers took up their positions, 20, 30, 40, . . . 80 feet, along the line indicated. A Table of tests, consisting of ten series of ten number-words each, was compiled.[1] These series were given at the various points along the line, and the subjects wrote down the numerical symbols of the words so far as they were heard. Results were computed on the basis of percentages of accuracy in each series, part credit being given where syllables, and not whole words, were heard correctly. There were ten observers. They are indicated in the Tables by the letters A, B, ... J.[2]

[1] These words (the original lists devised by the writer) were as follows:

I. 6, 42, 87, 34, 29, 61, 94, 18, 70, 53; II. 92, 33, 12, 75, 5, 27, 8, 69, 47, 50; III. 19, 28, 35, 60, 51, 86, 3, 98, 79, 45; IV. 96, 20, 41, 15, 89, 72, 57, 38, 2, 65; V. 13, 25, 46, 32, 7, 83, 68, 54, 73, 90; VI. 31, 4, 23, 14, 48, 67, 85, 91, 52, 76; VII. 37, 26, 49, 8, 66, 17, 74, 58, 93, 82; VIII. 81, 22, 13, 6, 30, 78, 64, 97, 44, 55; IX. 62, 43, 39, 84, 14, 56, 7, 99, 21, 77; X. 3, 40, 95, 36, 16, 59, 10, 63, 24, 71.

Like the revised lists given in the paper, these series are comparable as regards the relative frequency of low, medium, and high consonants (Wolf's classes). They are not exactly equivalent, however, in point of primary number-words present in each. After using them in the first part of the experiments, revised series (p.36), which met precisely this second criterion of equivalence, were substituted.

[2] The subjects, whose assistance the author takes this occasion of acknowledging, were Professor I. M. Bentley and Dr. J. W. Baird, of the Department of Psychology; Messrs. R. B. Waugh and W. A. Frayer, who had had some training as observers in psychological work; and Messrs. R. W. Palmer, L. E. Palmer, T. E. Faxon, G. L. Genung, C. L. Rand, and R. G. Marvin.

Table I gives the average (percentages) of accuracy for each individual at the various ranges at which he was tested. Each percentage is the average of accuracy for all the series used with a given individual and for a single distance. The figure in parenthesis after the percentages indicates the number of series on which the average is based.

TABLE I.

Range in feet.	20	25	30	40	50	60	65	70	75	80
Person.										
A.			75.0(3)	60.0(3)	51.6(3)		33.0(6)		32.5(2)	38.3(3)
B.	90.0(1)	95.0(1)	79.2(5)	73.7(4)	68.7(3)	61.1(1)	28.2(5)	55.4(3)		47.5(6)
C.	90.0(1)		66.2(4)	56.2(6)	66.2(1)	43.7(1)	35.0(6)	40.0(1)	38.7(2)	29.6(4)
D.			62.5(5)	56.1(5)	47.5(5)	52.5(2)	47.5(2)	42.5(2)		
E.	95.0(1)	72.5(1)	71.2(2)	47.5(3)	36.4(5)	46.2(1)	53.7(2)	42.5(1)		37.9(3)
F.	97.5(2)	55.0(1)	83.7(1)	65.0(1)	66.2(1)	66.2(1)		47.5(1)		41.2(1)
G.			45.2(5)	38.1(5)	46.6(6)	40.0(2)	33.7(2)			
H.	65.0(1)	55.0(2)	48.7(1)	52.5(1)	60.0(1)	37.5(1)		37.5(1)		42.5(1)
I.	83.7(3)		65.0(1)	60.0(1)	86.2(1)	27.5(1)		30.0(1)		32.5(1)
J. (deaf ear.)	23.7(1)		15.4(3)	2.5(1)	10.0(1)	5.0(1)		0.0(1)		0.0(1)
Totals. (A-I)	87.36	66.5	66.14	55.00	53.30	46.70	35.50	45.1	35.5	37.30
M. V.	10.2	21.2	12.7	9.4	11.9	10.0	10.1	7.6	4.3	8.2

This Table indicates, from the standpoint of individual results, that while there is a general decrease in audibility as the distance increases, yet the decrease is neither constant nor uniform. For example, the percentages for A and B show a fairly uniform decrease through the middle ranges; but there is an unexpected rise in the figures at the 80-foot range. This inequality is due, apparently, to reflection of sound-waves from the farther end of the room—a factor which makes results at the longer ranges of this particular room difficult of interpretation. The results indicate, further, that not only are the changes in intensity of sound not uniform as the range increases, but that over a considerable middle part of the range the intensity of the sounds may be approximately constant, so that differences in distance do not give corresponding differences in intensity. See, for example, the percentages of C and G at 30, 40 and 50 feet; or of E at 40-65 feet.

We are aware of the possibility of error in these conclusions due to the small number of cases considered; particularly, since the test-series used in this part of the experimental work lack perfect equivalence. Whatever influence might arise from either of these sources, however, would scarcely be felt in the

averages for observers A—I, which are given at the bottom of the Table.

The lack of uniformity, both in the individual and the general averages, suggests strongly the limitations—not to say the inadequacy—of the commonly accepted method of speech tests—the method of extreme range of accurate hearing. It shows that, under the given conditions, there is no simple relation between length of range and the goodness of hearing. Over and above these general results, the tests just considered throw some light upon the relative ease of perception of the various elementary number-words which make up our lists. Table II shows the percentage of accurate audition for each syllable at four of the ranges. The last column gives the percentage of right cases (sounds heard correctly) for each syllable at all distances. The total number of syllables recorded was twenty-five hundred.

TABLE II.

RANGE	30 Feet.		50 Feet.		65 Feet.		80 Feet.		All ranges.	
	WORD	% R.	WORD	% R.	WORD	% R.	WORD	% R.	WORD	% R.
1st	twen	96.4	six	87.9	six	86.0	six	92.3	six	90.4
2nd	six	94.9	-ty	80.0	seven	71.1	-ty	66.6	-ty	76.7
3rd	-ty	89.4	eight	77.5	-ty	64.9	two	61.9	seven	70.2
4th	eight	79.4	twen	71.4	two	57.1	seven	57.1	twen	68.0
5th	seven	79.2	seven	70.7	twen	54.5	twen	36.8	two	64.6
6th	nine	68.9	two	70.3	eight	41.3	three	25.0	eight	58.7
7th	two	66.6	-teen	48.1	nine	22.5	eight	22.8	-teen	38.6
8th	one	65.0	one	42.8	five	13.0	-teen	22.2	nine	37.8
9th	thir	58.8	three	40.9	three	12.5	four	19.0	one	34.2
10th	-teen	57.6	thir	38.8	-teen	11.7	five	16.29	thir	32.4
11th	five	47.8	five	38.6	four	6.9	thir	13.6	five	31.8
12th	four	45.6	nine	32.0	one	6.6	one	11.7	three	25.8
13th	three	25.9	four	19.6	thir	0.1	nine	9.6	four	22.8

In this Table, "one" is heard moderately well throughout the shorter ranges and poorly at the longer. "Two" holds a fairly constant position, at 60%-70%. "Four," which has a medium audibility at thirty feet, drops at the longer distances. "Five" holds its medium place up to fifty feet, but falls later. "Six," on the other hand, maintains a high place throughout. Similarly, "Seven" keeps a high position, though it drops off a little at the longest range.

One may say, in general, on the basis of the Table, that (1) there are distinct differences in the ease of audition for the different number-syllables; although (2) these differences are not constant throughout the various ranges. It is obvious, *e. g.*, that 'six' and 'ty' are better heard, at any one of the distances used, than 'four' and 'five;' but it is also obvious that the

advantage which one syllable has at, say thirty feet, may be transferred to another syllable at sixty-five or eighty feet.

This rise and fall of syllables in the Table is significant. If change in distance at which a sound is heard means simply alteration in intensity, then all words should suffer the same fate, as the range is increased; but if the phonetic elements undergo change by reason of re-enforcement and absorption, then we should expect to find precisely what, as a matter of fact, we do find in our results. We find that 'twen,' *e. g.*, falls from ninety-six to thirty-seven percentage of right cases where 'two' falls only from sixty-seven to sixty-two. The one syllable suffers rapid dissolution; the other is heard practically as well at eighty feet as at thirty. Other similar comparisons may be drawn.

We may say, in concluding this section, that the distribution of 'right' cases — whether we consider gross averages of series (Table I), or averages for individual phonetic elements (Table II) — shows that the physical conditions under which the experiments were made were too complex and too variable to give a definite and unambiguous measure of auditory acuity in terms of distance. Any test which, like ours and like speech tests in general, involves reflecting and absorbing surfaces complicates hopelessly both the absolute intensity and the qualitative integrity of the sounds employed and, therefore, violates conditions which we found at the beginning of this section to be indispensable to the success of the method.

II. *Method of Degree of Accuracy.* The root of the difficulty in the method just criticised lies in its attempting to measure audition in terms of a 'normal' distance where no 'normal' distance can be secured. It should, however, be possible to save a remnant of the traditional method by making — not distances[1] — but the number of right cases at *a given constant distance* the basis of computation. This modification we have attempted to carry through in the experiments now to be described.

The second group of experiments was performed, for the most part, in two suites of rooms in the psychological laboratory, at Cornell University. One suite, here designated "A," consisted of a lecture room 24 by 42 feet, adjoining which, by a door near one corner, was a small room 12 by 22 feet; the subjects were seated in the corner of the room diagonally opposite the door, and the range was laid off along the diagonal. The experimenter occupied three constant positions, two on the diagonal, giving a 20-foot and a 40-foot range, respectively,

[1] Various distances are used in the following experiments, but only by way of finding the most convenient conditions under which to work.

and a third position, ten feet away from the diagonal and within the smaller adjoining room. This gave three constant ranges of 20, 40, and 50 feet, in room A. The second suite, "B," was quite similar except in size; the large room was 42 by 18 feet, and the small room 8 by 20 feet. The range extended along the median line of the large room to a 20-foot and a 40-foot point, and from the latter at right angles through an open doorway to a 55-foot point in the smaller room. Room "B," had, accordingly, three constant ranges, 20, 40, and 55 feet. The two rooms, "A" and "B," from their size and the different location of the ranges, may be regarded as distinctly unlike as regards acoustical conditions.

The method consisted in submitting to the subjects a series of ten words at 20 feet, a second series at 40 feet, a third at 50 feet, and so on until all ten lists were given, distributed in each case as follows: three lists at 20 feet, three lists at 40 feet, and four lists at 50 or 55 feet (according as the tests were in room "A" or "B"). Only one ear at a time was used.

TABLE III.

			20 Feet.		40 Feet.		50 Feet in A. 55 Feet in B.	
Obs.	Room	Ear		M. V.		M. V.		M. V.
K	A	L	80.5	9.2	85.7	5.3	46.2	13.3
		R	66.1	2.7	83.3	2.2	43.3	5.0
	B	L	91.6	3.9	90.2	3.5	54.3	7.4
		R	91.1	6.3	83.3	5.0	66.6	8.3
			82.3	9.0	85.6	2.3	52.6	7.8
L	A	L	86.9	8.3	86.6	8.9	65.1	10.2
		R	86.1	2.6	90.0	6.7	56.6	9.1
	B	L	87.2	9.2	92.4	3.3	61.8	9.3
		R	94.4	5.2	95.5	2.9	73.3	8.3
			88.6	2.8	91.1	2.8	64.2	5.0
M	A	L	72.2	4.1	82.2	5.9	60.0	4.2
	B	L	90.0	4.4	85.0	5.5	57.9	5.4
			81.1	8.9	83.6	1.4	58.9	1.0

Table III includes results from two sets of experiments carried out in rooms "A" and "B." Reading the tables horizontally, one gets the percentage of correct cases for series of ten words each,[1] at a given range. The letters A and B in the

[1] Most of the results are based upon twenty series for each line; a few upon ten series.

Table designate the rooms used for the particular series of tests; K, L, and M, the observers.[1]

Assuming an acoustical difference in the two rooms, "A" and "B," we might expect, if our method is sufficiently delicate, to find some indication of it in comparing the results of similar tests given in the two rooms. Now the results do show higher average, range for range, for both the right and left ear, in room B than in room A. There are only two places in which B gives slightly lower averages than A, and there the range is 55 feet as against 50 feet in A. There is, accordingly, practical uniformity of higher averages for B than for A. This can only be accounted for as due to better acoustical conditions in room B. But it is fair to maintain that since the proposed method is precise enough to reflect in its results such acoustical differences, its delicacy, in actual tests of auditory capacity, is assured.[2]

But the most important issue of the method appears when we compare observer with observer within the same room and at the same range. Here we come to the real test of the method. The following Table gives the order, from worst to best, of the three observers and over the three ranges.

TABLE IV.

	Room	20	40	50 and 55
Left Ear	A	M, K, L	M, K, L	K, M, L
	B	L, M, K	M, K, L	K, M, L
Right Ear	A	K, L	K, L	K, L
	B	K, L	K, L	K, L

We find in the Table that, at the fifty and fifty-five foot ranges, the order for the observers is invariable. With both ears and in both rooms L hears with greatest and K with least acuteness. At the smaller distances, the order is not entirely constant. In considering the results, it must be borne in mind that no one of the observers is noticeably hard of hearing. We cannot expect, therefore, to discover large differences of capacity. It must be noted, further, that the small distances give percentages too near one hundred to be very significant. It is only where acuteness is considerably reduced—as it is in the last column—that we could expect to look for distinct and uniform differences. Where the method is used with marked im-

[1] These included the following Cornell students whose help is acknowledged with thanks: R. P. Butler, A. B. Truman, and O. Goehle.

[2] The fact that the forty-foot averages are higher than the twenty-foot is in striking confirmation of our contention regarding the effect of reflection upon intensity.

pairment of auditory capacity, we may expect still more striking and unambiguous results. To make sure of this, the writer tested three individuals, N, O and P,[1] whose respective ages were nineteen, thirty-five and sixty years and whose hearing was plainly different.

The experiments, which were carried out in two adjacent rooms in a dwelling, gave as an average of five series of ten words each and with a range of eighteen feet, the following percentages of right cases:

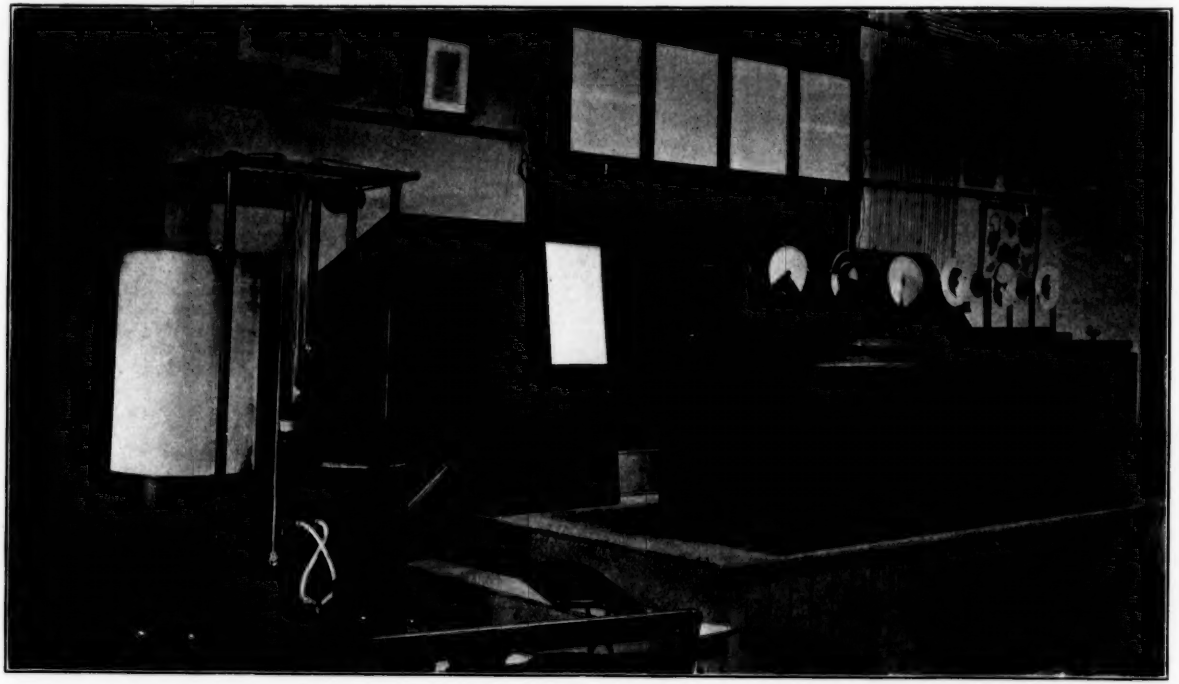
$N = 100 \pm 0$; $O = 80 \pm 7.3$; $P = 51 \pm 3$.

Even with so limited a number of results, we are safe in recommending the method of Degree of Accuracy. To be sure, it gives values which are valid only under the particular conditions of the experiments; but this is true, as well, of the method of maximal range which we have been obliged to discard as untrustworthy. Our own method, besides laying claim to vastly greater accuracy than the old method, is much more easily carried through, since it can be used in rooms of ordinary size,—the only necessity being sufficient space to give a fairly great number of incorrect answers. When once the proper range has been found, all other ranges should be discarded and the percentages reckoned on the basis of a single distance.[2]

---

[1] J. A., G. and C. T. Andrews.

[2] Experiments are under way in the Cornell Laboratory on a speech-method which will, if successful, eliminate entirely the factor of distance and provide a standard series of tests which can be employed without regard to local conditions.



SOME NEW APPARATUS.

E. B. TITCHENER.

In the last number of this *Journal* I described some apparatus for demonstration and class experiments.[1] Certain pieces were not completed at the time when that paper was written; so that I was not able to give illustrations. My present purpose is to make good this deficiency.

The accompanying half-tone plate shows four large apparatus, to which reference was made in my former paper.

1. *The new model Whipple gasometer.* The double gasometer for acoustic and general laboratory work was described and figured by Dr. Whipple in this *Journal*, XIV, 1903, 107 ff. A new model of the instrument is shown to the left of the accompanying plate. It has been rendered more compact, by the substitution of a metal for the original wooden frame. It has been mounted on a wheeled platform, so that it can be moved from room to room of the laboratory. Outlet cocks have been soldered into each tank, near the floor, so that the water can be drawn off and the apparatus correspondingly lightened for moving. The butterfly valves are encased, not in wooden boxes, but in felt-lined brass tubes. These can readily be taken apart if anything goes wrong with the flaps. Minor improvements have been introduced as regards the lateral guidance of the moving tanks, the overhead travelling of the chain, and the raising and shifting of the counterweight.

There is thus no change in principle, though the apparatus as a whole is a good deal more serviceable than it was in its first form. I may add that our inner stationary tanks have shown a tendency to rust into holes. We have learned wisdom by experience, and now keep some borax in solution in the contained water. The mending of a tank is, at best, an awkward business, and means the dismantling of at least half the apparatus.

2. *The frame for adaptation and after-images.* This is seen standing on the platform, to the right of the gasometer. The plate shows the frame in readiness for an experiment. The black-white sheet drops down, as was explained in my former paper, behind the black front, and exposes a gray sheet upon which the complementary after-images are seen.

[1] Commemorative Number, 175 ff; XIV, 1903, 439 ff.

3. *The contrast frame.* This is shown in place, upon the wall behind the lecturer's desk. The four colored papers are, from left to right, the Hering red, green, blue, and the Milton Bradley yellow. The gray strip is cut from Hering's gray paper, no. 11. The tissue paper front can be turned up, the colored papers and gray strip exposed, and the strip removed from the frame.

4. *The six-fold color mixer.* The mixer is also shown in place, upon the desk. The mechanism will be understood by reference to Fig. 1, which represents the parts at a quarter of their actual size.

a is the main frame of the mixer, made of cast iron. This is

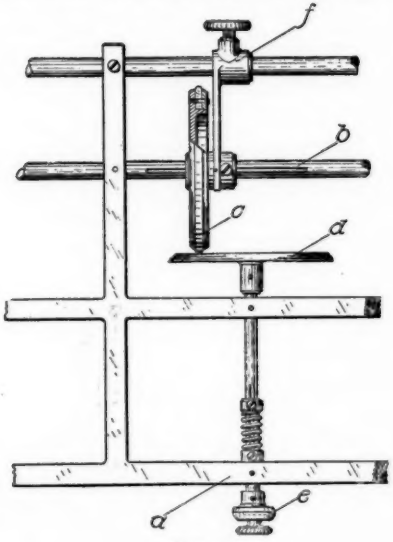


FIG. 1.

Detail of Color Mixer.

the cheapest material; and, with care, a good casting may be made. *b* is the main driving shaft, of wrought steel. It is driven by a motor, whose belt passes over the pulley shown at the extreme left of the instrument in the plate: the pulley may, of course, be placed at the center of the shaft, and the motor concealed beneath the lecturer's table, if desirable. The shaft carries *c*, the driving friction disc. This is made of alzinc, faced with leather. It actuates the driven friction disc, *d*, which is also of alzinc. It will be noticed that *d* tapers off at

the periphery, so that *c* can be brought on gradually, and metal strikes metal before the leather comes into action. As *d* turns, the spindle, carrying the clamp *e* for the colored papers, turns with it. *f* is the shifter, by means of which *c* can be set at any point upon the surface of *d*, or thrown off altogether. The remaining details of construction will be clear from the figure. It is evidently possible, while *b* is turning at a constant speed, to rotate the six color discs all at once at the same or at different speeds, or to rotate any one or more while the rest remain unmoved. The wooden base of the instrument stands upon a thick layer of felt, and the motor is set upon felt and boxed in: the noise may thus be reduced to a minimum. Even under the unfavorable conditions of my own lecture room — the desk acts as a resonance chamber, and the platform is built over a stairway and so acts as another — it is possible to continue lecturing, without strain of the voice, while the mixer is turning. The motor which we employ is a 1/6 H. P. Crocker-Wheeler; a less powerful motor would answer as well.

5. *The wall campimeter.* This is not shown in the plate. I mention it again because, by some fatality, the dimensions were wrongly given in my previous paper.[1] A strip of 2 m. in length, as seen from some parts of the lecture room, would serve only to take the stimulus a short distance beyond the macular region. The strip is, in reality, 12 m. long; the stimuli are discs, 20 cm. in diameter; there are three fixation crosses, two at the two ends and one at the center of the strip. The best results are obtained with a dark violet disc, which shows in the blue-yellow zone as a hard, staring blue, and a bright reddish brown, which shows in the same zone as a clear yellow.

The apparatus could easily be made permanent. I would suggest two vertical spring rollers, fastened to the wall at either end of the lecture room, and carrying each 6 m. of cloth; and a permanent wire, to be tightened by a ratchet, extending across the wall for the support of the upper edge of the cloth. The two strips of cloth could be hooked together where they meet, over the center of the blackboard; a draw-string along their lower edge would prevent curling. It would also be easy to devise a stimulus carriage, which should travel along wires from end to end of the cloth, and be operated from a single point by the lecturer. I doubt, however, from the partial estimates that I have so far made, whether the apparatus would be worth its cost. The manila strip, and the rod and disc in the hands of an itinerant lecturer, satisfy all requirements.

6. *A disc cutter.* I mentioned in my former paper that we

[1] *Ibid.*, 185; XIV, 449.

have used, to construct color discs, two discs of copper, turned accurately to the right size, and pierced at the center. The paper is laid between the copper discs, which are clamped together, and is trimmed round the edge with scissors.[1] I found at the beginning of the current academic year that the copper discs had suffered a good deal from student usage; their edges were flattened and dented. It therefore seemed worth while to make a different and less damageable apparatus for disc cutting. We now employ the instrument shown, quarter size, in Fig. 2.

e is a circular base of hard wood. Upon it is screwed an aluminium disc, *a*, of the same size. The base is pierced at the

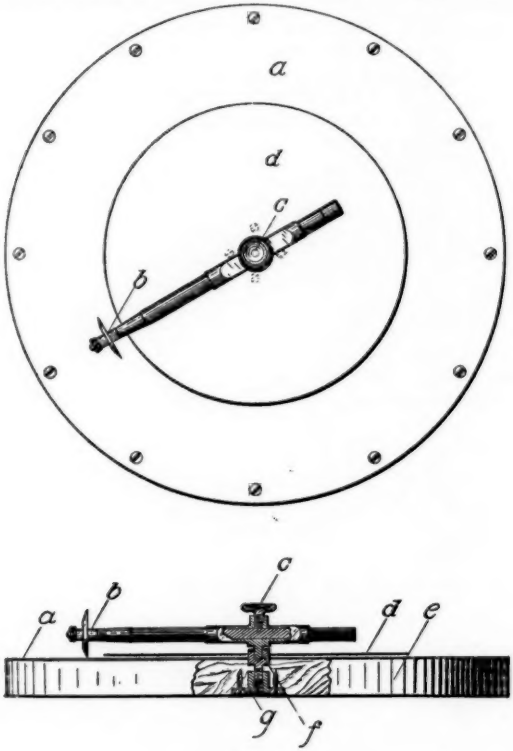


FIG. 2.
Disc Cutter.

[1] *Ibid.*, 182; XIV, 446.

center by a screw post, *f*, which is held below by a split nut *g*. The top of the screw post is hollowed to receive a boss from the cross-arm which carries the knife, *b*, of tempered steel. The knife can be set for any required radius; the arm is held in place by the cup screw *c*. *d* is a loose disc of aluminium, fitting over the top of the screw post.

The top of the screw post is of the same diameter as the *large-sized* central holes of the Hering or Zimmermann discs. When a disc of this sort is to be cut, the paper is roughly shaped with scissors, and the central hole punched out. The paper is then laid on *a*; *d* is laid over the paper; the cross-arm is put in place, with *b* at the right distance from the center; and arm and knife are turned from *c*. The disc *d* serves to hold the paper flat for cutting. The boss that extends into the head of the screw post is of the same diameter as the *small-sized* central holes of the imported discs. When a disc of this sort is to be cut the paper is shaped and the center punched, as before. Then *f* is turned down flush with the surface of *a* (there is a slot in the head of the screw post not shown in the Fig., which takes the screw driver); the paper is laid over *a*; a small-holed *d* is laid over the paper; the boss passes down through *d* and paper into the head of *f*; and the cutting is done as before. It is thus necessary to have four *d*; two large, with large and small central holes, and two small, with similar holes. The instrument is cheap, strong and simple, and ought to last longer than the copper discs.

All the pieces here mentioned were made by Mr. F. A. Stevens, the mechanician of the Cornell laboratory. I am indebted to Mr. Stevens for many improvements of design and for suggestions as regards details of construction.

EBBINGHAUS' EXPLANATION OF BEATS.

I. M. Bentley and E. B. Titchener.

This paper grew out of a discussion between the writers of the theory of beats set forth by Ebbinghaus in his *Psychologie*. The discussion brought out the fact that certain aspects of the theory had been, if not differently understood, at least very differently weighted and estimated by two interested readers of the book. Further reading and discussion have led, first; to an agreement as to the main features of Ebbinghaus' exposition, and secondly to the following joint criticism.

The paper is confined specifically to the subject of its title. The writers desire to insist upon this point, as it constitutes their apology for the neglect of many articles, etc., reference to which might, in view of the general topic, be expected by the reader.

The Helmholtz Theory of Tonal Sensations.

According to the revised version of the Helmholtz theory, the basilar membrane may be considered as a system of stretched strings, in which every individual string moves in relative independence of all the others.[1] If a simple tone is struck outside the ear, its stimulus throws a part of the basilar membrane into sympathetic vibration. That fibre vibrates most strongly whose proper motion corresponds most nearly to the movement of the external stimulus. From this fibre the vibrations extend on both sides, with rapidly diminishing strength, to the adjacent parts of the membrane. The sympathetic vibration thus set up "is still sensible for the interval of a semitone."[2]

Helmholtz applies Johannes Müller's doctrine of specific sensory energies to the fibres of the cochlear nerve: difference of tonal pitch is correlated with difference in the nerve fibre excited.[3] If, then, we follow the strict letter of the theory, we must suppose that the basilar fibres which vibrate on either side of the principal fibre, in any given case of tonal stimulation, move at the rate imposed on them by the stimulus, *i. e.*,

[1] *On the Sensations of Tone*, trs. by A. J. Ellis, 1895, 146 *b*.
[2] *Ibid.*, 144 *c*, 146 *c*, 147 *d*. The German word for 'sensible' is *merklich*. The context is physiological, and the term means 'large enough, physically, to be worth taking account of.'
[3] *Ibid.*, 148 *b*, *c*.

at the same rate as the principal fibre, but nevertheless mediate their own sensations. We have a strip or bundle of basilar fibres, all vibrating at the same frequency — though with rapidly lessening intensity as we leave the center of the strip, and all exciting the central organ in their own specific ways.[1]

There is a further point dealt with by Helmholtz in an Appendix,[2] which becomes of importance later in our discussion. Theoretically, *i. e.*, if we make certain assumptions with regard to the structure of the basilar membrane and draw their mathematical consequences, "every excitement of a bundle of transverse fibres by the fundamental tone must be accompanied by weaker excitements of the unevenly numbered harmonic undertones." These undertone fibres excite their own nerve fibres, and should accordingly, mediate their own sensations. "There is nothing of the kind observable in the ear." The reason is that "the appendages of the basilar membrane probably greatly impede the formation of tones with nodes."[3]

The physiological substrate of a simple tonal sensation is, therefore, threefold. We have (1) the strong sympathetic vibration of the basilar fibre whose proper tone lies nearest the tone of the stimulus. We have (2) the weaker sympathetic vibration of adjacent fibres. We have (3) the weak sympathetic vibration, with formation of nodes, of fibres carrying the odd-numbered harmonic undertones. The fibres of the two latter groups ought to arouse each their own proper tone. The undertone fibres, however, are so quickly damped that they are not represented in sensation. Helmholtz does not seem to have faced the difficulty that arises from the sympathetic vibration of the fibres adjoining the principal fibre, — the difficulty that these fibres must, in strictness, arouse their own tones, while only one single tone (the tone of the principal fibre) exists for sensation.

It may be said, at once, that this difficulty is not an insuperable objection to the Helmholtz theory. It may be overcome in various ways. Thus Stumpf, of whose views we shall have more to say presently, thinks on physiological and psy-

[1] The easiest way to envisage the Helmholtz theory is to think of it as a one-to-one correlation of basilar fibre and specific nerve fibre. And there is nothing seriously misleading in such envisagement. It should, however, be remembered that Helmholtz himself speaks in terms, not of the basilar fibres, but of Corti's arches and Corti's fibres: see, *e. g.*, *ibid.*, 147 *b*, *c*. He nowhere commits himself to the statement that the number of basilar fibres is a measure of the number of possible tone sensations.

[2] *Ibid.*, 411 *b*, *c*.

[3] It appears (unfortunately we have not the earlier editions of the *Tonempfindungen* at hand) that these considerations were suggested to Helmholtz by Riemann's *Musikalische Logik*, 1877.

chological grounds that an accommodation of the specific energies of the nervous structures, within certain narrow limits of stimulation, is not only possible but even highly probable.[1]

Ebbinghaus' Modifications.

Ebbinghaus accepts, without reserve, the fundamental idea of the Helmholtz theory.[2] In order, however, to meet certain difficulties of application, he proposes two modifications. (1) The first is directed against an extreme interpretation of the doctrine of specific energies. The cells of the cochlear nerves, or their central representatives, are so constituted as to mediate tonal sensations under the influence of rhythmical stimuli. Originally, Ebbinghaus says, any cell can mediate any sensation. But since the cell is connected with the basilar fibre, only a certain limited number of rhythms are, as a matter of fact, conducted to it. The rhythm that affects it most easily is, of course, the proper rhythm of its own basilar fibre; so that, by a process of physiological adaptation (*Gewöhnung*), it comes to respond most quickly and most strongly to this particular form of stimulus. It can, however, also respond, within certain limits, to other rhythms.[3]

Suppose, then, that a simple tonal stimulus is given. It sets in vibration a strip of the basilar membrane, a group of adjacent basilar fibres. These fibres vibrate, one and all, in the rhythm of the incoming stimulus. The corresponding cells mediate, not each its own sensation, but all the same sensation, *i. e.*, the sensation directly correlated with the stimulus. They do this because they still retain something of their primitive indifference of function: they are not so specifically adapted that they cannot take up rhythmical impulses which vary, in some small degree, from the rhythm of their proper fibre. The cell that is most nearly attuned to the stimulus arouses the strongest excitation; the neighboring cells repeat the same excitation, only in weaker form.

(2) Helmholtz had remarked that the fibres carrying certain harmonic undertones must vibrate, with the formation of nodes, on the operation of a given tonal stimulus. He supposed these fibres to mediate their own sensations; and accounted for the fact that their tones are not heard by the weight of the structures attached to the basilar membrane.

---

[1] *Tonpsychologie*, ii, 1890, 111 ff., with other references there given. At the time that the *Tonpsychologie* was written, Stumpf was an adherent of the Helmholtz theory. He has now given it up: see, *e. g.*, *Beitr. z. Akustik u. Musikwiss*, i, 1898, 18 *note*, 51 f. This fact should be borne in mind in reading the later divisions of the present paper.

[2] *Psychologie*, i, 1897, 316.

[3] *Ibid.*, 318 ff.

Ebbinghaus questions the soundness of this explanation; strings are capable of strong vibration even when very heavily loaded. He himself believes that the undertone fibres vibrate, with the formation of nodes. But he differs from Helmholtz in supposing that not the odd-numbered only, but all the undertone fibres, down to a certain limit, are involved; and, further, that the partial vibrations as such constitute the stimulus to the relatively indifferent cells. The cell is primarily adapted to the rhythm of its proper fibre. It is secondarily adapted, as we have seen, to the rhythms of adjacent fibres. It is also secondarily adapted to rhythms that are twice, three times, four times, etc., as rapid as the rhythm of its proper fibre. The undertone fibres consequently mediate precisely the same sensation as do the fibres of the strip set in direct vibration by the stimulus.[1]

An illustration will make the matter clear. Suppose that a tone of 600 vs. in the 1 sec. is sounded outside the ear. The effect upon the cochlea is (*a*) that the 600-fibre and its cell are thrown into strong vibration. At the same time (*b*) the neighboring fibres and cells, above and below, are thrown into weaker vibration, and so mediate the same excitation in weaker form. Further, (*c*) the 300-fibre vibrates at twice its natural rate, the 200-fibre at three times its natural rate, and so on. The cell of the 300-fibre accepts, takes up, the 600-rhythm, and mediates (if we may use the expression) the 600-sensation. The cell of the 200-fibre also accepts the 600-rhythm, and mediates, less strongly, the same sensation. These long fibres mediate the sensation proper to the shorter 600-fibre by reason of the fact that their cells are fairly well accustomed to vibrations two, three, etc., times as rapid as those normal to the fibres upon which they sit.

We have nothing to object to these modifications. It must only be remembered that Ebbinghaus makes them under stress of the psychological facts, and that they are both entirely hypothetical. Helmholtz gives one conjectural interpretation of the function of the cochlea; Ebbinghaus gives another. If, now, it is a question of cutting a physiological theory to suit the needs of psychology, we shall incline to accept the theory, whosesoever it may be, that without doing violence to the established facts of physiology best takes account of the results of introspection. On the whole, Ebbinghaus' 'adaptation' seems to us to offer a better working hypothesis than Helmholtz' 'specific energy.' But both theories alike contain too much of guess-work and analogical reasoning to allow us to strike a decision between them. The test, for the psychologist,

[1] *Ibid.*, 320.

lies just in their power to explain. The test applied to Ebbinghaus' theory in the present paper is its explanation of the phenomena of beats.

The Phenomena of Beats.

There can be no doubt that Helmholtz regarded beats, primarily, as *intensive* fluctuations of tonal sensation.[1] There can be no doubt, either, that he began his investigation with the slow beats formed when the two primary tones lie very near together upon the scale, and worked from these to the more rapid beats produced when the tones lie farther apart. Hence, beats appeared to him to be a very simple phenomenon, to be explained, as simply, by reference to the interference of vibrations. "Corti's arches,"[2] he says, "may be made to beat by two tones sufficiently near in pitch to set the same Corti's arches in sympathetic vibration at the same time. If, then, . . . the intensity of auditory sensation in the nerve fibres involved increases and decreases with the intensity of the elastic vibrations, the strength of the sensation must also increase and diminish in the same degree as the vibrations of the corresponding elastic appendages of the nerves."[3]

If we grant the introspective assumptions, this explanation is valid. Two neighboring tones are sounded. Each sets in vibration its own strip of the basilar membrane. These strips overlap; the same basilar fibres are affected by both stimuli, and there is interference. We therefore hear precisely what we should hear if there were no beats, except that the mass o sound is interrupted (or is subject to intensive fluctuation) at intervals corresponding to the difference in the vibration frequencies of the primary tones. The explanation is straightforward, and follows directly from the theory. The initial difficulty of the theory, the difficulty that a vibrating strip of the basilar membrane ought to give us a medley of different tonal sensations, of course remains.

Unfortunately, however, the facts as reported by introspection are less simple than Helmholtz supposed. We take Stumpf's account.

"(*a*) If I give two tones, about a semitone apart, in the middle region of the scale (*e. g.*, g^1-sharp and a^1 on the violin), I hear the two primary tones, but also, over and above these, a third tone which lies between them, somewhat nearer the lower than the higher. This third tone has a very soft coloring, and with keen attention is localized within the ear; it is this tone which beats, while the primary tones remain con-

[1] *Sensations of Tone*, 164 *c*, 165 *c*, *d*, 169 *a*. Stumpf, *Tonpsychologie*, ii, 450; *Beitr.*, 1, 4, *note*.

[2] Used in the sense explained *ibid.*, 147 *b*.

[3] *Ibid.*, 166 *c*, 167 *b*, *c*, 172 *a*, *b*.

stant. The two primary tones are, in my judgment, noticeably weakened, — more than is customary when two tones are sounded at the same time.

"(*b*) If I take tones that lie farther apart, in the same region of the scale (*e. g.*, g^1 and a^1), I do not hear any middle tone, but only the two primaries; and these two seem themselves to beat. If, however, I turn the attention more particularly to the one of them, this always appears to be the beating tone.

"(*c*) If, on the other hand, I take two tones that lie much nearer together than a musical semitone, so that they approximate the difference limen for simultaneous tones, I get one tone, and that beating. It is difficult to say whether it lies between the primaries." [1]

Stumpf's Explanations.

Stumpf seeks to account for these three observations in terms of the Helmholtz theory as modified by the hypothesis of physiological accommodation (accommodation of neighboring nerve fibres or their ganglion cells). If we grant Stumpf's presuppositions, the explanations are satisfactory. They are, in brief, as follows.

It is clear that Obs. (*c*)[2] is the typical observation required by the Helmholtz theory, and agrees with the observations made by Helmholtz himself. We need, therefore, spend no time upon it.

Obs. (*a*) demands more discussion. Stumpf supposes, with Helmholtz, that the two basilar strips, the g^1-sharp strip and the a^1 strip, partially overlap. We might, therefore, expect a priori to hear both the g^1-sharp and the a^1, and to hear both beating. What we actually hear is an intermediate beating tone, and the primaries sounding, smoothly but weakly, on either side. To explain the introspective phenomena, Stumpf appeals to his principle of nervous accommodation. There will be one intermediate nerve fibre, affected at equal strength by both forms of stimulation, for which the twofold "Erregungsweise am kräftigsten eintreten muss." It is then a simple extension of the principle of accommodation of nervous energies (*Bildung einer Gesammtenergie*) to assume that this intermediate nerve fibre or ganglion cell, giving its specific response under intensive stimulation, will constrain the neighboring fibres or cells in the direction of its own specific energy: so that the nervous structures excited fall, for purposes of sensation, into three groups, — an upper and a lower, replying specifically to the primary stimulations, and an intermediate, replying specifically to the intensive stimulation produced by

[1] Tonpsychologie, ii, 480 ff. [2] *Ibid.*, 488 f.

the joint operation of the two primary stimuli. In other words, we hear an intermediate tone, which carries the beats, and the two outlying tones, not beating, and weakened by the withdrawal (so to say) of certain of their constituent nerve fibres or nerve cells to the specific functional activity of the central zone.[1]

If we grant the presuppositions, this explanation is entirely acceptable. Stumpf gives further and very ingenious explanations of the peculiar coloring of the third tone, its localization, its pitch, and so on. We need not here enter into these details.

We come to Obs. (*b*). In this case, there is, according to Stumpf, a very slight overlapping of the basilar strips. The double excitation is thus too weak to arouse the intermediate nerve fibres or cells to their specific response. They beat; but they beat in the tone of the singly excited fibre or cell group above or below them. We therefore hear only the two primary tones, sounding more strongly than in Obs. (*a*); and we hear them beating, for the reason that a certain portion of their physiological substrates is affected by both stimuli.[2] The shift of the beating tone with shift of attention is readily intelligible.

In fine, then, Stumpf is able to give, in terms of the Helmholtz theory and his own hypothesis of nervous accommodation, a coherent and adequate explanation of the three observations. His explanation of Obs. (*a*) is especially happy. It is, in effect, a compromise—if one may speak of a compromise before both of the extremes have been formulated—between Helmholtz' doctrine of specific energies, and Ebbinghaus' doctrine of the original functional indifference of the cochlear cells.[3]

Ebbinghaus' Explanations.

Ebbinghaus accepts Stumpf's description of the phenomena. After describing Obs. (*b*) and (*c*), he proceeds to Obs. (*a*), as follows. "Nur bei geringer Höhendifferenz der Töne, etwa bis zu einem Halbton, verhält es sich, worauf Stumpf erst aufmerksam machte, etwas anders. Man hört dann einen *zwischenliegenden* Ton als Träger der Schwebungen, während die beiden objektiv erzeugten Töne entweder abgeschwächt und ruhigbleibend daneben hörbar sind oder auch—bei den geringsten Höhendifferenzen—ganz verschwinden. Eine Erklärung dieser *Zwischentöne* oder *resultierenden Töne* (Melde) wird später gegeben werden."[4]

Turning to the passage referred to, we read as follows:

---

[1] *Ibid.*, 484 ff.

[2] *Ibid.*, 487 f.

[3] Nevertheless, as will presently appear, we have a possible physical explanation of the *Zwischenton*, acceptance of which would render the accommodation hypothesis unnecessary.

[4] *Psychologie*, i, 1897, 304 f.

"Diese [Schwebungen] sollen [nach der Helmholtz'schen Theorie] dadurch zustande kommen, dass bei dem gleichzeitigen Erklingen zweier hinreichend benachbarter Töne die mitschwingenden Gebiete der Basilarmembran nicht mehr ganz getrennt bleiben, sondern teilweise übereinandergreifen, und die Bewegungen nun an den mittleren Fasern miteinander interferieren. Das passt vortrefflich, wenn die beiden objektiven Töne sehr nahe bei einander liegen. Dann hört man, wie oben mitgeteilt, in der That den nach der Theorie zu erwartenden *Zwischenton* als Träger der Schwebungen."[1]

Could anything be less satisfactory? The *Zwischenton* is precisely what the Helmholtz theory does *not* explain; it is precisely what we should *not* expect from that theory. What the Helmholtz theory would give us, in its strict form, is a medley of tones, some of which (the intermediate tones) are beating. What it under no circumstances can give us is the actual fact of observation,—two steady primaries and one beating intermediate tone.[2]

We may add that this Obs. is not either explained by Ebbinghaus' own modification of the Helmholtz theory. If the nervous appendages of the two basilar strips are 'adapted' to take up the two primary rhythms over their whole extent, there is no reason whatsoever to expect the appearance of an intermediate tone. This tone cannot, of course, be due to the sympathetic vibration of the undertone fibres, since these give the same sensation as the primary strips.[3]

As regards Obs. (*a*), therefore, Ebbinghaus leaves us just where we were left by Helmholtz. It is as if Stumpf had never written.

We pass to Obs. (*b*). "Aber es [the Helmholtz explanation] passt nicht mehr," Ebbinghaus goes on, "bei grösseren Entfernungen der primären Töne, etwa von einem Halbton ab.

---

[1] *Ibid.*, 317.

[2] Were it not for Helmholtz' specific energies, his theory might easily be so extended as to cover the intermediate tone of Obs. (*a*); for under certain conditions the superposition of two pendular vibrations of nearly the same frequency gives rise, theoretically, to a resultant vibration which lies midway between. See Rayleigh; *Theory of Sound*, 2 ed., i, 1894, 49,71; ii, 1896, 443, 450. In confirmation of these theoretical deductions, R. H. M. Bosanquet (*Phil. Mag.*, xi, 5 ser., 1881, 421) actually heard both the intermediate tone of Obs. (*a*) and the beating mass of Obs. (*c*) as lying midway between the primaries. Now Helmholtz' theory of analysis makes the beats of Obs. (*c*) mark the limit of Ohm's law; *i. e.*, beats arise when analysis is replaced by interference. The modification offered by the evidence just cited simply sets the limit of the law one degree higher, and brings not only the beating mass of (*c*) but also the intermediate tone of (*a*) under interference.

[3] Having broken away from H.'s specific energies, E. might well have explained the *Zwischenton* on the physical basis of the preceding note. But he rests on H.'s explanation, which is plainly inadequate.

Dann verschwindet der Zwischenton vollkommen, die Schwebungen aber, die noch an ihm haften sollen, verschwinden nicht auch, sondern bleiben hörbar noch bei beträchtlich grösseren Intervallen, und zwar haften sie deutlich an den beiden primären Tönen. Als Träger der Schwebungen können mithin, wie die Sache auch oben bereits dargestellt wurde,[1] keine anderen Fasern der Basilarmembran in Betracht kommen als die den schwebenden Tönen direkt zugeordneten. Und nur wenn diese bis auf einen Halbton und weniger aneinander rücken, kommt es zu einem für die Empfindung merklich werdenden Uebereinandergreifen der beiden Erregungsgebiete."[2]

The vehicles of the beats are, in this case, not the intermediate overlapping fibres, but the fibres directly correlated with the stimulus rhythms. How do these fibres mediate beats? By virtue of their inelasticity. "Die mitschwingenden feinen Teilchen der Basilarmembran . . . haben doch zweifellos nur geringe elastische Kräfte, und können ihre Bewegung beim Aufhören des objektiven Anstosses aus sich heraus keine nennenswerte Zeit weiter fortsetzen. . . . Jene Amplitudenschwankungen der objektiven Tonwelle werden demnach von den sämmtlichen mit ihr schwingenden Gebieten der Basilarmembran leidlich getreu mitgemacht. Erfolgen sie relativ langsam, so hören wir sie als Schwebungen, bei grösserer Schnelligkeit als Rasseln, bei noch grösserer als Rauhigkeit."[3]

Now no one can deny that the basilar fibres may be as inelastic as Ebbinghaus supposes. But, if they are, their inelasticity should be taken into account for the other observations.[4] For inelasticity comes to the same thing as interference, since both mean a failure of analysis. If the analysis is incomplete in the larger intervals of (*b*), it must also be incomplete in the smaller intervals of (*a*), and the 'resting primaries' remain a mystery. In other words, E.'s appeal to inelasticity advances the limit of auditory analysis to the most remote separation of the primaries that gives rise to beats (*b*). The appeal must

[1] See p. 303.

[2] *Ibid.*, 317.

[3] *Ibid.*, 1902, 322; *cf.* 1897, 303. On p. 322 occurs the sentence: "Die mitschwingenden feinen Teilchen der Basilarmembran dürfen nun nicht etwa gedacht werden wie stimmgabelartige Gebilde, wie es seitens der Helmholtzschen Theorie eigentlich geschieht." How this can be said to 'eigentlich geschehen' on the part of the Helmholtz theory, in view of Helmholtz' own discussion of the damping of vibrations in the ear (*Sensations of Tone*, 142 *b* to 143 *d*), we fail to see.

[4] We have sought in vain, in Ebbinghaus' pages, for any definite qualification or limitation of this attribute of inelasticity. If it extends as far as beats extend, then we might surely look for fluctuations in the difference tones,—since they, like beats, depend upon an intensively fluctuating stimulus. *Cf.* also M. Meyer's remarks upon certain mistuned intervals. *Zeits.*, xvi, 1897, 15; Stumpf's *Beitr.*, ii, 1898, 39.

then *a fortiori* cover all lesser separations, and the smooth-sounding tones of (*a*) become anomalous.

Again, Ebbinghaus has, in terms of his own modification of the Helmholtz theory, an explanation of Obs. (*b*). The basilar strips directly correlated with the stimuli are so wide apart that there is no overlapping. But the undertone fibres are also vibrating; and the interval between these fibres gets steadily smaller and smaller, as the fibres themselves become longer. A point must be reached, before very long, at which the undertone strips will overlap, and at which there will consequently be beats due to Helmholtzian interference. There is, thus, no necessity for the appeal to inelasticity. The beats of Obs. (*b*) may be due to phenomena of interference among the undertone fibres. Ebbinghaus himself makes no mention of this possibility.

Our criticism of Ebbinghaus, then, amounts to this: that he has multiplied causes for Obs. (*b*), and has altogether failed to account for the intermediate tone and the resting primaries of Obs. (*a*). We do not say how far interference is due to inelasticity; we do not know. But if the fibres of the basilar membrane are inelastic, there is no place for the smoothly sounding tones of the latter Obs. —

Although the matter of beats is far from settled, the writers venture to suggest that the case is not altogether hopeless. In regard to the three introspective observations, they offer the following considerations. Obs. (*c*), the beating mass, is probably correlated with cochlear vibrations which repeat with more or less fidelity the resultant from two pendular vibration rates. Obs. (*a*) depends, in part, upon two analyzed components; and in part, upon an unanalyzed residue possessing an intermediate frequency. Obs. (*b*) is, finally, to be referred to an almost complete analysis, in the sense of Ohm's law. The beating of the primaries may be explained by a small area of interference,—an area insufficient to mediate a distinct tone. The argument for Obs. (*b*) is supported by three bits of evidence. First, in Stumpf's introspections, it was difficult to localize the beat: the two primaries '*seem* themselves to beat;' but if either is attended to, that one '*always appears* to be the beating tone.' Secondly, weak and circumscribed interference might well give rise to intensive fluctuations without, however, producing a distinct tone (*cf.* weak partials discoverable only by beats). Thirdly, in our own experience, tones (*e. g.*, g^1 and a^1) which, when weak, are heard beating separately, give when excited more intensively a suggestion of an additional intermediate tone.

THE PROOF AND MEASUREMENT OF ASSOCIATION BETWEEN TWO THINGS.

By C. SPEARMAN.

CONTENTS.

	PAGE
Introductory.	72
Part I. Elementary correlation, and "accidental deviations."	73
1. Requirements of a good method.	73
2. Standard methods explained.	77
3. Comparison by rank.	79
4. Auxiliary methods.	82
Part II. Correction of "systematic deviations."	88
1. Systematic deviations generally.	88
2. "Attenuation" by errors.	89
3. Limits of associative problems.	92
4. "Constriction" and "Dilation."	94
5. "Distortion."	95
6. Criticism of prevalent working methods.	96
7. Number of cases desirable for an experiment.	99

INTRODUCTORY.

All knowledge—beyond that of bare isolated occurrence—deals with uniformities. Of the latter, some few have a claim to be considered absolute, such as mathematical implications and mechanical laws. But the vast majority are only *partial;* medicine does not teach that smallpox is inevitably escaped by vaccination, but that it is so generally; biology has not shown that all animals require organic food, but that nearly all do so; in daily life, a dark sky is no proof that it will rain, but merely a warning; even in morality, the sole categorical imperative alleged by Kant was the sinfulness of telling a lie, and few thinkers since have admitted so much as this to be valid universally. In psychology, more perhaps than in any other science, it is hard to find absolutely inflexible coincidences; occasionally, indeed, there appear uniformities sufficiently regular to be practically treated as laws, but infinitely the greater part of the observations hitherto recorded concern only more or less pronounced *tendencies* of one event or attribute to accompany another.

Under these circumstances, one might well have expected that the evidential evaluation and precise mensuration of tendencies had long been the subject of exhaustive investigation

and now formed one of the earliest sections in a beginner's psychological course. Instead, we find only a general naïve ignorance that there is anything about it requiring to be learnt. One after another, laborious series of experiments are executed and published with the purpose of demonstrating some connection between two events, wherein the otherwise learned psychologist reveals that his art of proving and measuring correspondence has not advanced beyond that of lay persons. The consequence has been that the significance of the experiments is not at all rightly understood, nor have any definite facts been elicited that may be either confirmed or refuted.

The present article is a commencement at attempting to remedy this deficiency of scientific correlation. With this view, it will be strictly confined to the needs of practical workers, and all theoretical mathematical demonstrations will be omitted; it may, however, be said that the relations stated have already received a large amount of empirical verification. Great thanks are due from me to Professor Haussdorff and to Dr. G. Lipps, each of whom have supplied a useful theorem in polynomial probability; the former has also very kindly given valuable advice concerning the proof of the important formulæ for elimination of "systematic deviations."

At the same time, and for the same reason, the meaning and working of the various formulæ have been explained sufficiently, it is hoped, to render them readily usable even by those whose knowledge of mathematics is elementary. The fundamental procedure is accompanied by simple imaginary examples, while the more advanced parts are illustrated by cases that have actually occurred in my personal experience. For more abundant and positive exemplification, the reader is requested to refer to the under cited research,[1] which is entirely built upon the principles and mathematical relations here laid down.

In conclusion, the general value of the methodics recommended is emphasized by a brief criticism of the best correlational work hitherto made public, and also the important question is discussed as to the number of "cases" required for an experimental series.

PART I.

Elementary Correlation and "Accidental Deviation."

1. Requirements of a Good Method of Correlation.

(a) Quantitative expression.

The most fundamental requisite is to be able to measure our observed correspondence by a plain numerical symbol. There

[1] 'General Intelligence,' determined and measured, to appear in a subsequent number of this *Journal*.

is no reason whatever to be satisfied either with vague generalities such as "large," "medium," "small," or, on the other hand, with complicated tables and compilations.

The first person to see the possibility of this immense advance seems to have been Galton, who, in 1886, writes: "the length of the arm is said to be correlated with that of the leg, because a person with a long arm has usually a long leg and conversely."[1] He then proceeds to devise the required symbol in such a way that it conveniently ranges from 1, for perfect correspondence, to 0 for entire independence, and on again to −1 for perfect correspondence inversely. By this means, correlations became comparable with other ones found either in different objects or by different observers; they were at last capable of leading to further conclusions, speculative and practical; in a word, they now assumed a scientific character.

Mathematically, it is clear that innumerable other systems of values are equally conceivable, similarly ranging from 1 to 0. One such, for instance, has been worked out and extensively used by myself (see pp. 86 ff). It therefore becomes necessary to discuss their relative merits.

(b) The significance of the quantity.

Galton's particular system is defined and most advantageously distinguished from all the others by the important property, that if any number of arms, for instance, be collected which are all any amount, $x\sigma_a$, above mean, then the corresponding legs will average $rx\sigma_l$ above the mean (with a middle or "quartile" deviation[2] of $\sigma_l \sqrt{1-r^2}$); where σ_a = the quartile variation of the arms, σ_l = that of the legs, and r is the measure of the correlation.

But another—theoretically far more valuable—property may conceivably attach to one among the possible systems of values expressing the correlation; this is, that a measure might be afforded of the *hidden underlying cause of the variations.* Suppose, for example, that A and B both derive their money from variable dividends and each gets $1/x^{th}$ of his total from some source common to both of them. Then evidently their respective incomes will have a certain tendency to rise and fall simultaneously; this correspondence will in any of the possible systems of values always be some function $1/x$, but in only one of them will it actually be itself $= 1/x$; in such a favored case, if A and B get, say, 20% of their respective incomes from the common source, the correlation between these two incomes will also show itself as 0.20; and conversely, if A's income happens to be found correlated with that of B by 0.20, then

[1] "Proceedings Royal Society of London," Vols. XL and XLV.
[2] Commonly, but misleadingly, termed the "probable error."

there is a likelihood that 0.20 of A's income coincides with 0.20 of B, leaving to either 0.80 disposable independently. The observed correlation thus becomes the direct expression of the relative amount of underlying influences tending for and against the correspondence.

In the above imagined instance, this desirable expressiveness belongs to the same above system of values proposed by Galton (and elaborated by Pearson). But this instance is exceptional and fundamentally different from the normal type. Evidently, A and B need not necessarily derive exactly the same proportion of their incomes from the common source; A might get his 0.20 while B got some totally different share; in which case, it will be found that the correlation is always the geometrical mean between the two shares. Let B be induced to put *all* his income into the common fund, then A need only put in $0.20^2 = 0.04$, to maintain the same correlation as before; since the geometrical mean between 0.04 and 1 is equal to 0.20.

Now, in psychological, as in most other actual correspondences, A and B are not to be regarded as in the fixed bisection of our first case, but rather as in the labile inter-accommodation of our second case. Hence A, in order to be correlated with B by $1/x$, must be considered to have only devoted $1/x^2$ (instead of $1/x$) of his arrangement to this purpose, and therefore to still have for further arrangements $1 - 1/x$,[2] which will enable an independent correlation to arise of $\sqrt{1 - 1/x^2}$. In short, not Galton's measure of correlation, but the *square thereof*, indicates the relative influence of the factors in A tending towards any observed correspondence as compared with the remaining components of A tending in other directions.

(c) Accuracy.

From this plurality of possible systems of values for the measure of the correlation must be carefully distinguished the variety of ways of calculating any one of them. These latter, again, have various advantages and disadvantages, of which the principal is their respective degrees of liability to "accidental deviation."

For, though the correlation between two series of data is an absolute mathematical fact, yet its whole real value lies in our being able to assume a likelihood of further cases taking a similar direction; we want to consider our results as a truly representative *sample*. Any one at all accustomed to original investigation must be aware how frequently phenomena will group themselves in such a manner as to convincingly suggest the existence of some law—when still more prolonged experiment reveals that the observed uniformity was due to pure hazard and has no tendency whatever to further repeat itself.

Luckily, this one great source of fallacy can be adequately

eliminated, owing to the fact that such accidental deviations are different in every individual case (hence are often called the "variable errors") and occur quite impartially in every direction according to the known laws of probability. The consequence is that they eventually more or less completely *compensate one another*, and thus finally present an approximately true result. Such elimination, however, must always remain theoretically incomplete, since no amount of chance coincidence is absolutely impossible; but beyond certain limits it becomes so extremely unlikely that for practical purposes we can afford to neglect it. When a person loses 14 times running at pitch-and-toss, he can reckon that such a series would not occur by mere accident once in 9,999 times, and consequently he will feel justified in attributing the coincidence to some constant disturbing influence. Similarly, to estimate the evidential value of any other observed uniformity, we only require to know how nearly the odds against chance coincidence have approached to some such standard maximum as 9,999 to 1. But, as any standard must always be more or less arbitrary—some thinking it too lenient and others unnecessarily severe—it is usual to employ a formula giving not the maximum but the middle deviation or "probable error."[1] We may then easily find the probability of mere hazard from the following comparative table:

If the observed correlation divided by the probable error be =	1	2	3	4	5	6
then the frequency of occurence by mere hazard =	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{21}$	$\frac{1}{143}$	$\frac{1}{1250}$	$\frac{1}{19000}$

Now, the smallness of this probable error depends principally upon the number of cases observed, but also largely upon the mathematical method of correlation. Though a faultiness in the latter respect can theoretically be made good by increasing the range of the observations, yet such increase is not always possible, and, besides, has other grave disadvantages which will be discussed later on. Other things being equal, therefore, *the best method is that one which gives the least probable error.* For the benefit of the reader, this probable error should always be plainly stated; nothing more is required than a rough approximation; for while it is highly important to distinguish between a deduction worth, say, 0.9999 of perfect certainty and one worth only 0.75, it would be a mere splitting of straws to care whether a particular experiment works out to a validity of 0.84 or to one of 0.85.

(d) Ease of application.

[1] In the proper use of this expression.

The most accurate ways of calculation are generally somewhat difficult and slow to apply; often, too, there occur circumstances under which they cannot be used at all. Hence, in addition to a standard method, which must be used for finally establishing the principal results, there is urgent need, also, of *auxiliary methods* capable of being employed under the most varied conditions and with the utmost facility.

But here a word of warning appears not out of place. For such auxiliary methods are very numerous and their results, owing to accidents, will diverge to some extent from one another; so that the unwary, "self-suggested" experimenter may often be led unconsciously—but none the less unfairly—to pick out the one most favorable for his particular point, and thereby confer upon his work an unequivocality to which it is by no means entitled. Any departures from the recognized standard methods are only legitimate, either when absolutely necessary, or for mere preliminary work, or for indicating comparatively unimportant relations.

2. Standard Methods Explained.

(a) Correlations between variables that can be measured quantitatively.

This may be regarded as the normal type of correlation. Its standard method of calculation is that discovered by Bravais,[1] in 1846, and shown by Pearson, in 1896,[2] to be the best possible. Pearson terms this method that of "product moments."

The formula appears most conveniently expressed as follows:

$$r = \frac{Sxy}{\sqrt{Sx^2 . Sy^2}}$$

where x and y are the deviations of any pair of characteristics from their respective medians,

xy is the product of the above two values for any single individual,

Sxy is the sum of such products for all the individuals,

Sx^2 is the sum of the squares of all the various values of x,

Sy^2 is similarly for y,

and r is the required correlation.

A simple example may make this method clearer. Suppose that it was desired to correlate acuteness of sight with that of hearing, and that for this purpose five persons were tested as to the greatest distance at which they could read and hear a standard alphabet and sound respectively. Suppose the results to be :

[1] "Memoires par divers savants," T, IX, Paris, pp 255-332.
[2] "Phil. Trans., R. S., London," Vol. CLXXXVII, A, p. 164.

Person	Sight	Hearing
A	6 ft.	6 ft.
B	7	11
C	9 (median)	12
D	11	10 (median)
E	14	8

then, we get

	x	y	xy	x^2	y^2
A	−3	−4	+12	9	16
B	−2	+1	−2	4	1
C	0	+2	0	0	4
D	+2	0	0	4	0
E	+5	−2	−10	25	4

$$Sxy = \begin{cases} +12 \\ -12 \end{cases} \quad Sx^2 = 42 \quad Sy^2 = 25$$

so that $r = \dfrac{+12-12}{\sqrt{42 \times 25}} = 0$, and there, thus, is no correspondence, direct or inverse.

The "probable error" between any obtained correlation and the really existing correspondence has been determined by Pearson, as being "with sufficient accuracy" when a fairly large number of cases have been taken,

$$= 0.674506 \frac{1-r^2}{\sqrt{n\ (1+r^2)}}$$

For discussion of correlation between characteristics whose distribution differs considerably from the normal probability curve as regards either "range" or "skewness," reference may be made to the works below.[1] It may be remarked that the method of "product moments" is valid, whether or not the distribution follow the normal law of frequency, so long as the "regression" is linear.

(b) Correlation between characteristics that can *not* be measured quantitatively.

In the example quoted by Galton, of correspondence between the length of arm and that of leg, it may be noted that the correspondence is proportional quantitatively; a long arm has a tendency to be accompanied by a leg not only long, but long to the same degree. Now, in many cases, such proportionality is by the nature of things excluded; a printed word is possibly remembered better than one heard; but, nevertheless, we cannot, in accordance with the preceding formula, ascertain whether degrees of visuality are correlated to retentiveness of memory, seeing that in the former case there do not exist any degrees, a word being simply either seen or not seen. Perhaps even

[1] Udny Yule: "Proc. R. S. London," Vol. LX, p. 477.
Pearson: "Phil. Trans. R. S. London," Vol. CLXXXV, I A, p. 71; Vol. CLXXXVI, I A, p. 343, and Vol. CXCI A, p. 229.
G. Lipps: "Die Theorie der Collectivgegenstände," Wundt's Phil. Stud., Vol. XVII.

more numerous are those cases where proportionality does indeed exist, but practically will not admit of being measured; for instance, it is probable that conscientiousness is to some extent a hereditary quality, yet we cannot well directly determine whether brothers tend to possess precisely the same amount of it, owing to the fact that we cannot exactly measure it.

In all such cases we must confine ourselves to counting the frequencies of coexistence. We can easily find out how often seen and spoken words are respectively remembered and forgotten. It has proved quite feasible to divide the children of a school generally into "conscientious" and "non-conscientious," and then to measure how much brothers tend to be in the same division; when we have proved this simple association, we may provisionally assume correlation of quantity also; that is to say, if the "conscientious," generally speaking, have a particular degree of tendency to possess brothers likewise "conscientious," then boys with excessively tender scruples will have the same degree of tendency to possess brothers with similarly excessive tenderness, while those with only a moderate amount of virtue will be thus correlated with brothers also of only moderate virtue; further, the ethical resemblance may be expected to repeat itself in cousins, etc., only reduced in proportion as the kinship is diminished.

For measurement of this non-proportional association, a standard method, which may be termed that of "cross multiples," has been elaborated by Sheppard,[1] Bramley-Moore, Filon, Lee, and Pearson. The formula is, unfortunately, too long and complicated to be usefully quoted in this place. It will be found in the under cited work,[2] together with its probable error as determined by Pearson.[3] In practice, it will generally have to be replaced by one of the more convenient methods to be next described.

3. Comparison by Rank.

This method of "cross multiples" is not only difficult and tedious of application, but also it gives a probable error nearly double that of "product moments."

Now, it can often be altogether escaped in the case of quantities not admitting absolute measurement, by substituting instead *comparison*. This other way will be discussed at some length, as it has been largely used by myself and is believed chiefly responsible for some successful experiments. All characteristics may be collated from two quite distinct aspects: either (as in example of visual and auditory acuteness) by actual

[1] "Phil. Trans.," Vol. CXCII, A, p. 141.
[2] "Phil. Trans.," Vol. CXCV A, pp. 2-7.
[3] "Phil. Trans.," Vol. CXCV A, 10-14.

mensuration, or else by order of merit; we might say that a student, A, obtained 8,000 marks in an examination, while B only got 6,000; or, instead, we might say that A was third out of 100 candidates, while B was only 20th. Precisely the same method of calculation may be again used in the latter case, simply substituting the inverse ranks, 97, 80, etc., for the performances, 8,000, 6,000, etc.

(a) Disadvantages of the "Rank" method.

In the first place, it may be objected that the observed correlation would then only hold good for persons of the same average difference from one another. For assuming, say, acute sight to be correlated with acute hearing; then the order of merit of A, B and C, as regards sight, is more likely to remain unaltered as regards hearing also, when the difference in their respective powers of vision is extremely marked, than when they are practically equal on the latter head. But the more numerous the persons experimented on, the less will be the average difference of faculty; it might, therefore, be supposed that the correlation would become continually less perfect as the experiments were made more extensive. This, however, would be a fallacy: 100 experimental subjects compared together by "Rank" would on the whole actually show appreciably the same average correlation as 1,000, provided, that in either case the subjects are selected by chance; the amount of the correlation is not really dependent upon the difference between the grades, but upon the relation of this difference to the mean deviation; and both of these increase together with the number of subjects. On the other hand, the correlation will undoubtedly diminish if the subjects be all chosen from a more homogeneous class; in a select training school for teachers, for example, general intelligence will throughout show smaller correlation with other qualities, than would be the case in a college for quite average young men of the same age; but this fact applies just as much to comparison by "Measurement."

The next possible objection is that comparison by rank bases itself upon an assumption that all the subjects differ from one another by the same amount, whereas A may differ from B five times as much as B differs from C. But such an assumption would only take place, if correspondence by rank were considered to be wholly equivalent to that by measurement; no such assumption is made; the two aspects are recognized to be theoretically distinct, but advantage is taken of the fact that they give correlational values sensibly equivalent in amount. Even against the small existing discrepancy may be set off a deviation of the same order of magnitude which is incurred when using measurement itself, owing to the practical necessity of throwing the cases into a number of groups.

The third and only solid objection is that rank affords a theoretically somewhat less full criterion of correspondence than does measurement; and the force, even of this argument, disappears on considering that the two methods give appreciably the same correlational values.

(b) Advantages of the "Rank" method.

The chief of these is the large reduction of the "accidental error." In a normal frequency curve, the outlying exceptional cases are much more spaced apart than are those nearer to the average; hence, any accident disturbing the position of these exceptional cases will have unduly great effect on the general result of the correlation; and owing to this inequality in the influence of the errors, the latter will not compensate one another with the same readiness as usual. Moreover, it is just these hyper-influential extreme cases where there is most likelihood of accidental errors and where there very frequently prevails a law quite different from that governing the great bulk of the cases. As regards the quantity of this gain by using rank (abstracting from the last mentioned point, which cannot well be estimated in any general manner) there should be no difficulty in calculating it mathematically. From a considerable amount of empirical evidence, the probable error when using the method of "product moments" with rank appears to become less than two-thirds of that given by the same method with measurement, and therefore only about one-third of that given by the method of "cross multiples."

The next advantage is that rank eliminates any disparity between the two characteristics compared, as regards their general system of distribution; such a disparity is often not intrinsic or in any way relevant, but merely an effect of the particular manner of gaining the measurement. By means of rank, a series presenting the normal frequency curve can be compared on even terms with another series whose curve is entirely different. This cannot well be done when using measurement. (See p. 78.)

Rank has also the useful property of allowing any two series to be easily and fairly combined into a third composite one.

(c) Conclusion.

From the practical point of view, it is so urgently desirable to obtain the smallest probable error with a given number of subjects, that the method of rank must often have the preference even when we are dealing with two series of measurements properly comparable with one another.

Theoretically, rank is at any rate preferable to such a hybrid and unmeaning correlation as that between essential measurements on the one side and mere arbitrary classification on the other. As the latter occur in most psychological correlations,

the only other resource would be to avoid measurements altogether by using the method of "cross multiples." But this trebles the size of the probable error, and therefore renders it necessary that the subjects should be no less than nine times as numerous; such an enormous increase, even if possible, would generally be accompanied by disadvantages infinitely outweighing the supposed theoretical superiority of method.

The above advantages are still further enhanced whenever dealing with *one-sided* frequency curves, such as are furnished by most mental tests. For in these cases the great bulk of influence upon the resulting correlation is derived exclusively from the very worst performances and is consequently of a specially doubtful validity.

In short, correlation by rank, in most cases a desirable procedure, is for short series quite indispensable, rendering them of equal evidential value to much longer ones treated by other ways. Luckily, it is precisely in short series that gradation by rank is practically attainable.

4. Auxiliary Methods.

These, as has been said, are only for use when there is adequate reason for not employing the above "standard" methods. Any number are devisable. Their resulting correlational values do not quite coincide with those found by the standard ways, but nearly enough so for most practical purposes.

(a) Auxiliary methods of Pearson.

Several very ingenious and convenient ones are furnished by him,[1] but all of similar type and requiring the same data as that of "cross-multiples."[2] They are therefore for use when the compared events do not admit of direct quantitative correlation. The following appears to combine facility and precision to the greatest degree:

$$r = \sin \frac{\pi}{2} \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}}$$

where the two compared series of characteristics, say P and Q are each divided into two (preferably about equal) classes; if the case is one where quantity exists but cannot be absolutely measured, P II will comprise the instances in which P is in manifest deficiency; but if the compared characteristics essentially exclude quantity, P II become the instances where P is

[1] "Phil. Trans. R. S. L.," Vol. CXCV, A, pp. 1 and 79.

[2] They are all refinements of the original formula, $r = \frac{ad - bc}{ad + bc}$, published by Yule, Proc. R. S. L.," Vol. LXVI, p. 23.

absent; similarly Q. Then,

a = the number of times that P I is accompanied by Q I
b " " " " " P II " " Q I
c " " " " " P I " " Q II
d " " " " " P II " " Q II.

If a + b is not very unequal to c + d, the probable error may be taken at about $1.1/\sqrt{n}$, where n = the number of instances in the whole of P or of Q.[1]

Returning to our previous illustration, suppose that it was desired positively to ascertain the merits of instruction by writing and by word of mouth respectively. Ten series, each consisting of ten printed words, have been successively shown to a class of twenty children, who each time had to write down by memory as many as they could. The experiment was next repeated, but reading the words aloud instead of showing them. Of the 2,000 visual impressions 900 were correctly remembered, while of the same number of auditory ones only 700 were retained.

Call the visual impressions P I
" " auditory " P II
" " remembered " Q I
" " forgotten " Q II

then a = 900, b = 700, c = 1,100, d = 1,300, and

$$r = \sin \frac{\pi}{2} \frac{\sqrt{900} \times \sqrt{1,300} - \sqrt{700} \times \sqrt{1,100}}{\sqrt{900} \quad \sqrt{1,300} + \sqrt{700} \quad \sqrt{1,100}} = 0.16$$

The probable error then comes to $1.1/\sqrt{4,000}$ = nearly 0.02, or about 1/8 of the above correlation; so that the latter would not occur by mere chance once in 100,000 times.

We thus see that there is at any rate good *prima facie* evidence of some superiority on the part of the visual sense. Also, if the experiment has been fairly executed and adequately described, any subseqent verification under sufficiently similar conditions, by other experimenters, should result in a concordant correlation, probably between 0.14 and 0.18, and certainly between 0.04 and 0.28.

Moreover, we have obtained a direct estimate of the importance of this apparent superiority of the visual sense; for the square of the correlation amounts to 0.025; so that of the various causes here tending to make the children remember some words better than others, the difference of sense impressed comes to about one fortieth part (see p. 75).

---

[1] More accurately, $\sin 0.1686\,\pi\,(1 - r^2)\,\sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$

(b) Method of proportional changes.

This is very often convenient, being especially applicable to a large number of psychological experiments, and so easy that the result can be approximately seen on inspection. Here,

$$r = \frac{3}{2}\frac{a - b}{a + b}$$

where a = the number of cases that have changed in accordance with the supposed correspondence, and b = the number that have changed in contradiction of it. The probable error again comes to $\frac{1.1}{\sqrt{n}}$.

Suppose, for example, we were demonstrating that intellectual fatigue may be satisfactorily investigated by the method of Griessbach.[2] With this view, we have applied his test to 100 boys before and after their lessons. In the latter case 68 of them have presented the expected duller sensitivity, but 32, on the contrary, have shown a finer discrimination than before work.

Now, clearly, had the correspondence been perfect, *all* the hundred would have become worse.[3] Thus,

$$r = \frac{3}{2}\frac{68 - 32}{100} = 0.54.$$

As the probable error comes to 0.11, our imaginary correlation is five times greater, and therefore would not have occurred by mere accident more than once in 1,250 times; so that we become practically certain that the sensitivity of the skin really does measure fatigue.

It now becomes easy to compare the quantity of this fatigue at different stages of work. Let us say that further experiments, after lessons lasting one hour longer than before, showed the correlation had risen to 0.77. Thereby we see that the influence of fatigue swells from 0.54^2 to $0.77,^2$ that is, from being 1/5 to being 3/5 of all the sources of variation in cutaneous sensitivity. Such a result has a very different scientific significance from, say, any conclusion that the average sensory threshold had enlarged by so many more millimetres.

---

[1] Hence, when the correlation is very complete, say, over 0.75, the above formula gives appreciably too large values; as the amount reaches 0.90 and 1, the first factor must be reduced from 3/2 to 5/4 and 1 respectively.

[2] This, as is well known, consists in determining the least distance apart at which two points of contact can be distinguished as being double and not single.

[3] Assuming, that is to say, that all the boys become fatigued by the lessons.

Moreover, our test can be *easily and precisely compared with any of the various other recommended procedures*, being more reliable than all which present smaller correlations and *vice versa*.

(c) Method of class averages.

It often happens that measurements (or ranks) are known, but not in such a way as to be able to use either the method of "product moments" or even any of the methods of Pearson. Under such circumstances, I have found it very useful to be able to apply the following relation:

$$r = \frac{d}{D}$$

More accurately, $r = \sin \frac{\pi}{2} \cdot \frac{a - b}{a + b}$

where d is the observed difference between the average measurement (or rank) of the P's accompanied by Q I and that of those accompanied by Q II, and D is the greatest difference that was possible (such as would have occurred, had the correspondence been perfect). If Q has been divided into two about equal portions, D will be equal to twice the middle or "quartile" deviation from the average in the whole series P; while if Q has been divided after the usual fashion into three such portions, only the two outer ones can be used and then $D = 2.87$ times the above middle deviation (again taken in the whole series P).

Suppose, for example, that we wish to ascertain whether the well known test of "reaction-time" gives any indication as to the person's general speed of movement. We try a hundred persons both in reaction-time and in speed of running 50 yards. Then we divide the reaction-time records into two classes, I containing all the quickest performers and II all the slowest. We now see how long these two classes of reacters took respectively to run the fifty yards, and what was the middle deviation from the average among all the runners taken together. Let us put the average of class I at 6 seconds, that of class II at 6.5 seconds, and the general middle deviation at 1.1 seconds. Then

$$r = \frac{6.5 - 6}{2 \times 1.1} = 0.23$$

The evidential value of the result is given approximately, even for small values of n, by the following relation:

$$\text{probable error} = \frac{1.17}{\sqrt{n}} \frac{\sqrt{n+1}}{\sqrt{n+2}}$$

where n is the *total* number of cases considered. In the three-

fold instead of twofold division, the probable error becomes nearly

$$\frac{1.4}{\sqrt{n}}\,\frac{\sqrt{n+1}}{\sqrt{n+2}}$$

In the above instance, we find that the observed correlation is little over double the probable error; as so much would turn up about once in six times by mere accident, the evidence is not at all conclusive. Therefore we must either observe many more cases — 600 would be necessary to reduce the probable error to 1/5th of the correlation — or else we must find a better method of calculation. If rank had been employed instead of measurement, the evidence would already have been fairly good, and could have been put beyond all reproach by the addition of another 150 observations. If rank had been employed in conjunction with the method of "product moments" or that of "rank differences," the required smallness of probable error could have been obtained by as few as 36 cases in all!

The method of "class averages" is especially valuable in deciphering the results of other investigators, where the average performances and the middle deviation are usually given (in good work), but not the data required for any of the other methods.

(d) Method of rank differences.

This method appears to deserve mention also, seeing that it seems to unite the facility of the auxiliary methods with a maximum accuracy like that given by "product moments." It depends upon noting how much each individual's rank in the one faculty differs from his rank in the other one; evidently, this will be nil when the correlation is perfect, and will increase as the correlation diminishes.[1]

[1] This general idea seems to have been first due to Binet and Henri ("La fatigue intellectuelle," p. 252-261), who, however, do not work it out far enough to obtain any definite measure of correlation. Accordingly, Binet makes little further attempt in later research (L'année psychologique, Vol. IV) to render it of service, and soon appears to have altogether dropped it (L'année psychologique, Vol. VI).

The same idea occurred to myself and was developed as above, without being at the time acquainted with the previous work in this direction by Binet and Henri. In obtaining the above formulæ I was greatly assisted by Dr. G. Lipps' showing generally that when an urn contains n balls numbered 1, 2, 3, . . , n, respectively; and when they are all drawn in turn (without being replaced); and when the difference is each time noted between the number on the ball and the order of its drawing; then the most probable (or middle) total sum of such differences, added together without regard to sign, will be $= \frac{n^2 - 1}{3}$.

Previously, I had only calculated this value for each particular size of n required by myself. Prof. Hausdorff further showed, generally, that such sum of differences will present a mean square deviation (from the above most probable value) $= \sqrt{\frac{(n+1)(2n^2+7)}{45}}$

The relation is as follows:

$$R = 1 - \frac{3Sd}{n^3 - 1}$$ [1]

where Sd is the sum of the differences of rank for all the individuals,

n is the total number of individuals,

and R is the required correlation.

The probable error will then be approximately, even for small values of n, $= 0.4/\sqrt{n}$.

To take again the example from p. 80, we number the five persons according to their order of merit in hearing and seeing respectively.

PERSON.	SEEING RANK.	HEARING RANK.	DIFFERENCE.
A	1	1	0
B	2	4	2
C	3	5	2
D	4	3	1
E	5	2	3
			Sd = 8

so that $$R = 1 - \frac{3 \times 8}{25 - 1} = 0,$$

and again we find that there is no correlation, direct or inverse.

This method, though very accurate and pre-eminently quick in application, has unfortunately four serious disadvantages.

It can be only used for ranks, and not immediately for measurements.

The probable error given is only that showing how great correlations may be expected from pure accident when there is no really existing correspondence between the two characteristics. It does not (like Pearson's probable error for the method of "product moments") directly show how much the observed correlation may be expected to differ by accident from any correspondence that does exist.

The various possible values of Sd are found to fall into a frequency curve of marked asymmetry; so that we cannot (as in all the other methods here given) take the *minus* values of R as representing so much *inverse* correlation. This defect could be remedied mathematically; but there are also other respects in which this side of the frequency curve appears unsuitable for our purpose, so that it is better to treat every cor-

[1] This formula becomes slightly incorrect, whenever two or more individuals are bracketed as having precisely the same rank; but the consequent error is usually too small to be worth considering.

relation as positive (which can always be done by, if necessary, inverting the order of one of the series).

Finally, this value R is not numerically equivalent to the "r" found by all the other methods, but for chance distributions appears $= \sqrt{r^3}$. So far, the proof of this relation is only empirical, but it rests on a large number of cases taken, however, only between 0.20 and 0.60. If it be accepted r can at once be found from the following table:

R	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1
r	0.13	0.22	0.34	0.44	0.54	0.63	0.71	0.79	0.86	0.93	1

PART II.

Correction of "Systematic Deviations."

1. Systematic Deviations Generally.

In the first part, we have seen that any correlational experiments, however extensive, can only be regarded as a "sample" out of the immense reality, and will consequently present a certain amount of accidental deviation from the real general tendency; we have further seen that this accidental deviation is measurable by the "probable error," whose determination, therefore, becomes an indispensable requisite to all serious research.

But now we are in danger of falling from Scylla into Charybdis. For after laboriously compiling sufficient cases and conscientiously determining the probable error, there exists a very human tendency to cease from labor and inwardly rejoice at having thus risen from common fallacious argument to the serene certainty of mathematics. But whether or not such complacency may be justifiable in pure statistical inquiry, it is at any rate altogether premature in the kind of research that we are at present contemplating; we are not dealing with statistics, but with a line of work so fundamentally different, that it may be aptly distinguished by the term of "statisticoids." Here the accidental deviation is not the sole one, nor even the most momentous; there are many other enemies who are unmoved by the most formidable array of figures. These consist in such deviations as, instead of merely being balanced imperfectly, lie wholly on the one side or the other. As in ordinary measurements, so too in correlation, we may speak, not only of "accidental," "variable," or "compensating" inaccuracies, but also of "systematic," "constant," or "non-compensating" ones.

These systematic deviations are of very varied nature, the most insidious being, as usual, *self-suggestion*. To take, for instance, one of our recent examples, suppose that we have applied the Griessbach test to a number of children before and

after their lessons, and have found the desired correlation between fatigue and cutaneous insensitivity, it still remains exceedingly difficult to convince ourselves that we executed our tests entirely without favor or affection; for it is almost impossible to determine a series of sensory thresholds without some general tendency, either to bring them towards the desired shape, or else—endeavoring to escape such bias—to force them in the opposite direction. To convince others of our impartiality may be harder still. Even this sort of deviation is to be remedied by our proposed exact method of procedure, for by it we obtain perfectly definite results which any impartial experimenters may positively corroborate or refute.

2. "Attenuation" by Errors.

From page 74 it will be obvious that a correlation does not simply depend on the amount of concording factors in the two compared series, but solely on the proportion between these concording elements on the one hand and the discording ones on the other. In our example, it did not matter whether A and B each had one pound or a thousand pounds in the common funds, but only whether the amount was a small or large fraction of their whole incomes. If the discordance, $1 - x$, be nil, then the concordance, x, is thereby perfect, that is, $= 1$; and if the influence of the discordant elements be sufficiently increased, then any concordance will eventually become infinitely small.

To consider a still more concrete example, suppose three balls to be rolled along a well-kept lawn; then the various distances they go will be almost perfectly correlated to the various forces with which they were impelled. But let these balls be cast with the same inequalities of force down a rough mountain side; then the respective distances eventually attained will have but faint correspondence to the respective original momenta.[1]

Thus it will be clear that here the accidental deviations have a new consequence simultaneous with, but quite distinct from, that discussed in the last chapter. For there, they impartially augmented and diminished the correlation, tending in a prolonged series to always more and more perfectly counterbalance one another; and in ordinary measurements, this is their sole result. But here in correlations, they also have this new effect which is always in the direction of "attenuating" the apparent correspondence and whose amount, depending solely on the size of the middle error, cannot be in the least eliminated by

[1] This fact has already been mathematically expressed in the last chapter by the value of correlation between two series being proportional (inversely) to the value of the middle deviations inside the series (see p. 86).

any prolongation of the series. The deviation has thus become general or "systematic."

Now, suppose that we wish to ascertain the correspondence between a series of values, p, and another series, q. By practical observation we evidently do not obtain the true objective values, p and q, but only approximations which we will call p′ and q′. Obviously, p′ is less closely connected with q′, than is p with q, for the first pair only correspond at all by the intermediation of the second pair; the real correspondence between p and q, shortly r_{pq}, has been "attenuated" into $r_{p'q'}$.

To ascertain the amount of this attenuation, and thereby discover the true correlation, it appears *necessary to make two or more independent series of observations of both p and q.* Then,

$$r_{pq} = \frac{r_{p'q'}}{\sqrt{r_{p'p'} \cdot r_{q'q'}}}$$

where $r_{p'q'}$ = the mean of the correlations between each series of values obtained for p with each series obtained for q.

$r_{p'p'}$ = the average correlation between one and another of these several independently obtained series of values for p.

$r_{q'q'}$ = the same as regards q.

and r_{pq} = the required real correlation between the true objective values of p and q.

Thus, if for each characteristic two such independent series of observations be made, say p_1 p_2 q_1 and q_2, then the true

$$r_{pq} = \frac{r_{p_1q_1} + r_{p_1q_2} + r_{p_2q_1} + r_{p_2q_2}}{4\sqrt{(r_{p_1p_2} \times r_{q_1q_2})}}$$

Should circumstances happen to render, say, p_1, much more accurate than p_2, then the correlations involving p_1 will be considerably greater than those involving p_2. In such case, the numerator of the above fraction must be formed by the geometrical instead of by the arithmetical mean; hereby the accidental errors of the respective observations cease to eliminate one another and therefore double their final influence; they also introduce an undue diminution of the fraction.[1]

In some exceptional and principally very theoretical cases, it may happen that either of the actual measurements, say p'_1 is

[1] By an inversion of the above formula, the correlation between two series of observations will be found a useful measure of the accuracy of the observations.

connected with q′ (or q) quite independently of p or any other link common to p'_2. Then, the correlation $r_{p'q'}$ will be to that extent increased without any proportional increase in $r_{p'p'}$; hence our above formula will fallaciously present too large a value.

A greater practical difficulty is that of obtaining two series sufficiently independent of one another. For many errors are likely to repeat themselves; even two separate observers are generally, to some extent, warped by the same influences; we are all imposed on by, not only the "Idola Specus," but also the "Idola Tribus" and the "Idola Fori." In such case, the above formula is still valid, only its correction does not go quite far enough,—a fallacy at any rate on the right side.

An actual instance will best show the urgent necessity of correcting this attenuation. In a correlation between two events, say P and Q, I obtained three independent observations both of P and of Q. The average correlation for those of P with those for Q was 0.38 ($= r_{p'q'}$); the average correlation of those for P with one another was 0.58 ($= r_{p'p'}$); the same for Q was 0.22 ($= r_{q'q'}$). Therefore, the correspondence between the real events, P and Q, comes by reckoning to $\frac{0.38}{\sqrt{0.58 \times 0.22}}$ = approximately 1; so that the correspondence, instead of being merely 0.38, appeared to be absolute and complete.

Attenuation by errors can also be corrected in another manner, which has the great advantage of an independent empirical basis, and therefore of not being subject to either of the two above mentioned fallacies besetting the other method. Hence, when the results coincide both ways, the fallacies in question may thereby be considered as disproved, for it is very unlikely that they should both be present and in such proportions as to exactly cancel one another. In this method, instead of directly employing the values p_1 p_2 p_3, etc., we amalgamate them into a single list; by this means we clearly eliminate *some portion* of the individual observational errors, and thereby we cause any really existing correspondence to reveal itself in greater completeness. Now, this increase in correlation from this partial elimination of errors will furnish a measure of the increase to be expected from an *entire* elimination of errors. Assuming the mean error to be inversely proportional both to this increase in the correlation and to the square root of the number of lists amalgamated, the relation will be:

$$r_{pq} = \frac{\sqrt[4]{mn}.r_{p''q''} - r_{p'q'}}{\sqrt[4]{mn} - 1}$$

where m and n = the number of independent gradings for p and q respectively,

$r_{p'q'}$ = the mean correlation between the various gradings for p and those for q,

and $r_{p''q''}$ = the correlation of the amalgamated series for p with the amalgamated series for q.

In the above quoted instance, the three observations for series P were amalgamated into a single list, and similarly those for series Q. Upon this being done, the two amalgamated lists now presented a correlation with one another of no less than 0.66 ($= r_{p''q''}$). Thus by this mode of reckoning, the real correspondence became

$$= \frac{\sqrt[4]{3\times 3}\times 0.66 - 0.38}{\sqrt[4]{3\times 3} - 1} = \text{once more approximately 1,}$$

so that this way also the correspondence advanced from 0.38 to absolute completeness.

If more than two independent series of observations are available, we may acquire additional evidence by trying the effect of *partial* amalgamation. Instead of throwing all our obtained values together, we may form a set of smaller combinations for each of the two compared characteristics, and then see the mean correlation between one set and the other. In our above instance instead of summarily considering $p'_1\ p'_2\ p'_3$, we can have $p'_1\ p'_2$, $p'_1\ p'_3$, and $p'_2\ p'_3$, and find out their mean correlation with similar values for q. This works out actually to 0.55. Hence

$$r_{pq} = \frac{\sqrt[4]{2\times 2}\times 0.55 - 0.38}{\sqrt[4]{2\times 2} - 1} = \text{approximately 1.}$$

Thus, again, by this third way, where both terms are the mean of 9 observed correlational values, the correspondence once more rises from the apparent 0.38 to the real 1.[1]

3. Limits of Associative Problems.

We have seen that "the length of the arm is said to be correlated with that of the leg, because a person with a long arm has usually a long leg and conversely;" also that this correlation is defined mathematically by any constant which deter-

[1] The exactness of the coincidence between the two methods of correction is in the above instance neither greater nor less than generally occurs in practice. It was singled out, in order to show that the formulæ still hold perfectly good even for such an enormous rise as from 0.38 to 1. The possibility of such a rise is due to the unusual conditions of the experiment in question, whereby the three observations of the same objective series presented the extraordinarily small inter-correlation of 0.22.

mines the function of any definite size of arm to the mean of the sizes of the corresponding legs. These terms, taken literally, are very wide reaching and express what we will call the "universal" correlation between the two organs.

But evidently not the most painstaking investigation can possibly secure any adequately representative sample for such universal correlations, even in the simple case of arms and legs. To begin with, they would have to be equally derived from every stage of growth, including all the prenatal period; since this is the most influential of all causes of variation in size. In the next place, they would have to come from every historical epoch, containing their fair proportion of big Cro-Magnons, little Furfoozers, etc. Further, they must impartially include every living race, from the great Patagonians to the diminutive M'Kabbas; also every social class, from the tall aristocrats to the under-sized slummers.

Practically, then, the universal correlation, even if desirable, is quite inaccessible. We are forced to successively introduce a large number of restrictions: the sample is confined to adults, to moderns, to some particular country, etc., etc. In a word, we are obliged to deal with a *special* correlation.

When we proceed to more narrowly consider these restrictions, it soon becomes clear that they are far from being really detrimental. For every serious investigation will be found to be directed, however vaguely and unconsciously, by some hypothesis as to the causes both of the correspondence and of the digression therefrom (see page 74). This hypothesis will determine a particular system of restrictions, such as to set the correspondence in the most significant relief.

But from these restrictions will at the same time proceed several kinds of grave errors. In the first place, since the restrictions are not explicitly recognized, they often are not carried out in a manner scientifically profitable; then, the result, however true, may nevertheless be trivial and unsuggestive. For instance, a series of experiments was recently executed by one of our best known psychologists and ended—to his apparent satisfaction—in showing that some children's school-order was largely correlated with their height, weight, and strength. As, however, no steps had been taken to exclude the variations due to difference of age, the only reasonable conclusion seemed to be that as children grow older they both get bigger and go up in the school! Such explanation turned out in fact to probably be the true and sufficient one.

The next fault to be feared is equivocality. For even if the controlling under-thought be good, yet its indistinctness in the mind of the experimenter causes the restriction to be carried out so unsystematically, that the results inevitably become ambiguous and fruitless.

The last is that, even with the clearest purpose, this specialization of the correlation is an exceedingly difficult matter to execute successfully. Only by a profound knowledge of the many factors involved, can we at all adequately exclude those irrelevant to our main intention.

Now, all such elements in a correlation as are foreign to the investigator's explicit or implicit purpose will, like the attenuating errors, constitute impurities in it and will quantitatively falsify its apparent amount. This will chiefly happen in two ways.

4. "Constriction" and "Dilation."

Any correlation of either of the considered characteristics will have been admitted irrelevantly, if it has supervened irrespectively of the original definition of the correspondence to be investigated. The variations are thereby illegitimately constrained to follow some irrelevant direction so that (as in the case of Attenuation) they no longer possess full amplitude of possible correlation in the investigated direction; the maximum instead of being 1 will be only a fraction, and all the lesser degrees of correspondence will be similarly affected; such a falsification may be called "constriction." Much more rarely, the converse or "dilation" will occur, by correlations being irrelevantly excluded. The disturbance is measurable by the following relation:

$$r_{pq} = \frac{r'_{pq}}{\sqrt{1 - r^2_{pv}}}$$

where r'_{pq} = the apparent correlation of p and q, the two variables to be compared,

r_{pv} = the correlation of one of the above variables with a third and irrelevantly admitted variable v.

and r_{pq} = the real correlation between p and q, after compensating for the illegitimate influence of v.

Should any further irrelevant correlation, say r_{pw}, be admitted, then

$$r_{pq} = \frac{r^1_{pq}}{\sqrt{1 - r^2_{pv} - r^2_{pw}}}$$

In the reverse case of "dilation,"

$$r_{pq} = r^1_{pq} \cdot \sqrt{1 - r^2_{pv} - r^2_{pw}}$$

These formulæ will be easily seen to be at once derivable from the relations stated on pages 74 and 75. Small, irrelevant

variations evidently do not affect the result in any sensible degree, while large ones are capable of revolutionizing it.

The following is an actual illustration of this constriction. I was investigating the correspondence between on the one hand intelligence at school lessons and on the other the faculty of discriminating musical pitch. The correlation proved to be 0.49. But, upon inquiry, it turned out that more than half of the children took lessons in music and therefore enjoyed artificial training as regards pitch; here, then, was a powerful cause of variation additional and quite irrelevant to the research, which dealt with the correspondence between the two natural faculties. When this disturbant had once been detected, there was no difficulty in eliminating its influence by the above formula; the correspondence between pitch discrimination and music lessons was measured at 0.61; so that the true required correlation became

$$\frac{0.49}{\sqrt{1 - 0.61^2}} = 0.62.$$

In this particular case, the more desirable course was open of eliminating the constriction, *practically*, by confining the experiment to those children who were learning music and therefore were on a sufficient equality as regards the training. The correlation then gained in this purely empirical way exactly coincided with the former result, being again 0.62.

5. "Distortion."

Whereas Attenuation and Constriction have wholly tended to reduce the apparent correlation, and Dilation to enlarge it, we now come to a third kind of impurity that may equally well reduce or enlarge. Its effect is thus analogous to the first consequence of accidental errors discussed in the first part of this article, but, unlike the latter, this Distortion does not in the least tend to eliminate itself in the longest series of observations.

Distortion occurs whenever the two series to be compared together both correspond to any appreciable degree with the *same* third irrelevant variant. In this case, the relation is given by

$$r_{pq} = \frac{r^1_{pq} - r_{pv} \cdot rqv}{\sqrt{(1 - r^2_{pv})\ (1 - r^2qv)}} \qquad (^1)$$

where r^1_{pq} = the apparent correlation between p and q, the two characteristics to be compared,

r_{pv} and r_{qv} = the correlations of p and q with some third and perturbing variable v,

[1] This same formula has already been arrived at, though along a very different route, by Yule. See Proc. R. S. L., Vol. LX.

and r_{pq} = the required real correlation between p and q, after compensating for the illegitimate influence of v.

Should the common correspondence with v have been irrelevantly excluded instead of admitted, the relation becomes

$$r_{pq} = r^1_{pq} \cdot \sqrt{(1 - r^2_{pv})\,(1 - r^2_{qv})} + r_{pv} \cdot r_{qv}$$

In the course of the same investigation above alluded to, but in another school, the correlation between school intelligence and discrimination of pitch turned out to be —0.25, so that apparently not the cleverer but the stupider children could discriminate best! But now it was observed that a superiority in discrimination had been shown by the older children, amounting to a correlation of 0.55; while, for a then unknown reason, the schoolmaster's estimate of intelligence had shown a very marked (though unconscious) partiality for the younger ones, amounting to a correlation of 0.65. Hence, the true correlation reckoned out to $\dfrac{-0.25 - 0.55 \times (-0.65)}{\sqrt{(1 - 0.55^2)\,(1 - [-0.65]^2)}}$ $= +0.17$. This latter low but direct correlation was — under the particular circumstances of the experiment — unquestionably about correct; so that the one originally observed of $= -0.25$ would have been entirely misleading.

6. Criticism of Prevalent Working Methods.

So far, our illustration of systematic deviation has been confined to instances taken from personal experience. But it might perhaps be thought that other workers avoid such perversions of fact by the simpler method of common sense. Unfortunately, such does not seem to have been at all the case; not once, to the best of my knowledge, has any partial association between two psychological events been determined in such a way as to present any good evidential value—these are strong terms, but, I think, hardly exaggerated.

Psychologists, with scarcely an exception, never seem to have become acquainted with the brilliant work being carried on since 1886 by the Galton-Pearson school. The consequence has been that they do not even attain to the first fundamental requisite of correlation, namely, a precise quantitative expression. Many have, indeed, taken great pains in the matter and have constructed arrays of complicated numerical tables; but when we succeed in orienting ourselves in the somewhat bewildering assemblage of figures, we generally find that they have omitted precisely the few facts which are essential, so that we cannot even work out the correlation for ourselves.

This lack of quantitative expression entails far more than merely diminished exactitude. For, in consequence, the ex-

perimenters have been unable to estimate their own results at all correctly; some have believed themselves to demonstrate an entire absence of correspondence, when the latter has really been quite considerable; whereas others have presented to the public as a high correlation what has really been very small and often well within the limits of mere accidental coincidence; these limits they have had no means of determining, and moreover their data were usually obtained in such a way as to make it unnecessarily large.

Seeing, thus, that even the elementary requirements of good correlational work described in the first part of this article have been so generally deficient, we cannot be surprised to find that the more advanced refinements of procedure discussed in the second part have been almost wholly unregarded; so that the final results are saturated and falsified with every description of impurity. In this respect, unfortunately, it is no longer possible to hold up even the Galton-Pearson school as a model to be imitated. The latter must now perform the very different office of saving us from detailed criticism of inferior work, by enabling us to form an opinion as to how far the defect permeates and vitiates even the best existent correlational research.

As example, we will take Pearson's chief line of investigation, Collateral Heredity, at that point where it comes into closest contact with our own topic, Psychology. Since 1898 he has, with government sanction and assistance, been collecting a vast number of data as to the amount of correspondence existing between brothers. A preliminary calculation, based in each case upon 800 to 1,000 pairs, led, in 1901, to the publication of the following momentous results:

COEFFICIENTS OF COLLATERAL HEREDITY.

Correlation of Pairs of Brothers.

PHYSICAL CHARACTERS. (Family Measurements.)		MENTAL CHARACTERS. (School Observations.)	
Stature	0.5107	Intelligence	0.4559
Forearm	0.4912	Vivacity	0.4702
Span	0.5494	Conscientiousness	0.5929
Eye-color	0.5169	Popularity	0.5044
(School Observations.)		Temper	0.5068
Cephalic index	0.4861	Self-consciousness	0.5915
Hair-color	0.5452	Shyness	0.5281
Health	0.5203		
Mean	0.5171	Mean	0.5214

Dealing with the means for physical and mental characters, we are forced to the perfectly definite conclusion, *that the*

mental characters in man are inherited in precisely the same manner as the physical.[1] Our mental and moral nature is, quite as much as our physical nature, the outcome of hereditary factors.

Now, let us consider how these coefficients of correlation will be affected by our "systematic deviations." To begin with, there is the "Attenuation" by errors; since it evidently cannot be assumed that the schoolmasters' judgments as to conscientiousness, temper, etc., are absolutely infallible. On page 90, it has been shown that deviation from this source may be estimated by the following formula:

$$r_{pq} = \frac{r_{p'q'}}{\sqrt{r_{p'p'} \cdot r_{q'q'}}}$$

To ascertain $r_{p'p'}$ and $r_{q'q'}$, I am aware of no precise data beyond that found in some experiments of my own, where the independent intellectual gradings for the same series of subjects correlated with one another on an average to the amount of 0.64. As on other occasions very competent persons have estimated this to be as much as should be expected, and as intelligence is about the most easily gradable of all the mental qualities mentioned by Pearson, there is so far no reason to suppose that his "great number of masters and mistresses" did on the whole any better. Hence, even *if* we could assume that the mistakes in estimating one brother were independent of the mistakes in estimating the other, then the true correlation would be about, not 0.5172, but $\frac{0.5172}{\sqrt{0.64 \times 0.64}} = 0.81$, an extent of difference that seriously modifies our impression of exactitude from all these coefficients to four places of decimals. When we further consider that each of these physical and mental characteristics will have quite a different amount of such error (in the former, this being probably quite insignificant), it is difficult to avoid the conclusion that the remarkable coincidence announced between physical and mental heredity can hardly be more than mere accidental coincidence.

Let us next proceed to irrelevant correlation, and take for our theme postnatal accidents connected on the one side with brotherhood and on the other with the mental qualities. Pearson's primary intention seems to have been to make his correlation as "universal" as possible, and in one place he expressly mentions that education is among the causes contributory to variation. Hence, he is no more than consistent, in that he forms his correlation without regard to the fact that the correspondence between the brothers' "conscientiousness," "popu-

[1]The italics are Pearson's.

larity," etc., must be in great measure due to their coming under the same home influences. But such a correlation can scarcely be accepted as scientifically valuable. For we do not really know anything precise about the assimilating effects of heredity, when our observed correspondence is perhaps chiefly due to the brothers having been similarly brought up — or even to such accidents as their being equally well dressed and having the same amount of hampers and pocket-money. Still less can we, then, fairly compare such a result with that obtained from physical measurements, where common home life has little or no effect. The factor of post-natal accidents, therefore, cannot but be regarded as *irrelevant*, and consequently the coefficients of correlation must be taken as hopelessly "distorted."

But even consistence cannot be upheld throughout the matter. For though the effect of post-natal life has thus been admitted with regard to education at home, it has perforce been excluded as regards public education. For only those brothers have been compared together who are at the same school; the coefficients of correlation would certainly diminish if those also could be included who are living in a totally different manner, have gone to sea, etc. The correlations are therefore also illegitimately "dilated."

If this work of Pearson has thus been singled out for criticism, it is certainly from no desire to undervalue it. The above and any other systematic errors are eventually capable of adequate elimination, and this article has itself, it is hoped, been of some use towards that purpose. Such correction will no doubt necessitate an immense amount of further investigation and labor, but in the end his results will acquire all their proper validity. My present object is only to guard against premature conclusions and to point out the urgent need of still further improving the existing methodics of correlational work, a method of investigation which he himself has so largely helped to create and by means of which he is carrying light into immense regions hitherto buried in the obscurity of irresponsible speculation. The fundamental difference between his procedure and that here recommended, is that he seeks large natural samples of any existing series sufficiently homogeneous to be treated mathematically; whereas here smaller samples are deemed sufficient, but they are required to be artificially selected, ordered, and corrected into full scientific significance. His methods are those of pure statistics; those inculcated here may be more aptly termed "statisticoids."

7. Number of Cases Desirable for an Experiment.

This leads us to the important question, as to how many cases it is advisable to collect for a single series of experiments. In actual practice, the greatest diversity has been apparent in

this respect; many have thought to sufficiently establish important correlations with less than ten experimental subjects, while others have thought it necessary to gather together at least over a thousand.

Now, a series of experiments is a very limited extract, whose disposition is, nevertheless, to be accepted as a fair sample of the whole immense remainder. Other things equal, then, the larger the sample, the greater its evidential value and the less chance of a mere occasional coincidence being mistaken for the permanent universal tendency.

This danger of accidental deviation has been discussed in the first part and there shown to be strictly measurable by the "probable error." We there saw, also, that this danger can never be entirely eliminated by *any sample however large*, so that it is necessary to accept some standard less rigorous than absolute certainty as sufficient for all practical pnrposes; usually, the danger of mere chance coincidence is considered to be inappreciable when a correlation is observed as much as five times greater than the probable error, seeing that mere chance would not produce this once in a thousand times. Hence, evidently, the accidental deviation depends, not only on the number of cases, but also on the largeness of the really existing correspondence; the more perfect the latter, the fewer the cases that will be required to demonstrate it conclusively; and this tendency is augmented by the fact that the probable error, besides varying inversely with "n," does so to a further extent with "r" (see formula). It was shown in the same part that the size of the probable error also varies according to the method of calculation—and to such an extent that twenty cases treated in one of the ways described furnish as much certitude as 180 in another more usual way. If the common trifold classification be adopted, an even greater number is required to effect the same purpose; and if the correlation be not calculated quantitatively at all, but instead be presented in the customary fashion to the reader's general impression, then no number of cases whatever appear sufficient to give reasonable guarantee of proof.

While thus the number of subjects is not by any means the sole factor in diminishing even the accidental deviation *it has no effect whatever upon the far more formidable systematic deviation*, except that it indirectly leads to an enormous augmentation thereof. When we are taking great pains to be able to show upon paper an imposing number of cases and a diminutive probable error, we are in the self same process most likely introducing a systematic deviation twenty times greater.

From all this, we may gather that the number of cases should be determined by the simple principle, that the measurements to be aggregated together should have their error brought to

the *same general order of magnitude*. An astronomical chronometer, with spring-detent escapement, is not the best travelling clock; nor is there any real advantage in graving upon a milestone (as has actually been done by an infatuated mathematician!) the distance to the nearest village in metres to three decimal places. Now, the present stage of Correlational Psychology is one of pioneering; and, instead of a few unwieldy experiments, we require a large number of small ones carefully carried out under varied and well considered conditions. At the same time, however, the probable error must be kept down to limits at any rate small enough for the particular object of investigation to be proved. For such a purpose a probable error may at present be admitted without much hesitation up to about 0.05; so that, by adopting the method of calculation recommended, two to three dozen subjects should be sufficient for most purposes. The precision can always be augmented subsequently, by carrying out similar experiments under similar conditions and then taking averages. Only after a long preliminary exploration of this rougher sort, shall we be in a position to effectually utilize experiments designed and executed from the very beginning on a vast scale.

PROFESSOR CATTELL'S STATISTICS OF AMERICAN PSYCHOLOGISTS.

By I. MADISON BENTLEY, Cornell University.

I should like to comment on certain points in Professor Cattell's recent census returns, published in the last number of this *Journal*. Criticism of the returns is difficult, both because they form a part of a more comprehensive report not yet published, and because the author's interpretation of the census figures is cautious and conservative. There are, however, two or three matters of interpretation upon which difference of opinion may, I believe, fairly be held.

(1) I am inclined to think that too strong emphasis is laid upon position in order of merit, as determined, in each case, by ten judgments. The interpretation seems to imply that the judges were individually capable of making only gross, not delicate, discriminations of merit. A complete concurrence, *e. g.*, is represented (position one) by a wide gap in the table. But this representation is just only if the gap reappears in the individual judgments. *Degree* of difference does not, however, seem to have been taken into account by the judges. This is simply an illustration. My point is, that more is got out of the average than the individual factors warrant. Another illustration of the same thing appears in the fact that the average assumes equal validity for all judgments. Now the admission is made that the judges did "not possess equal information or judgment" and, also, that "many of those who made the arrangement stated that they had but little confidence in its validity." I do not see how, under these circumstances, 'validity of judgment' can be measured, as the author proposes, by an m. v. of individual judgments from the average of the ten observations. The best informed and most judicious observer may have been just the one to vary most widely from the average.

(2) In grading men of so young a science as psychology, it should be noted that "contributions to the advancement" of the science (the meaning of 'merit' accepted in the statistics under discussion) depend not only upon age but upon length of service and upon the time at which reputation is acquired. M's contributions of twenty-five years may give him fourth place, let us say, while N's contributions, covering five years (perhaps at a different stage of the science), give him twelfth

place. It is, I take it, much more difficult to get a wide reputation in psychology now than it was ten or fifteen years ago. It seems unjust, at all events, to designate M— on the basis of his position in the table—a 'moderate genius,' and N, a 'man of talent.' Length of service and date of reputation may have been taken into account by Professor Cattell's judges; I see no evidence that they were.

(3) Similar allowance should be made for Table III, which gives geographical distribution of psychological students. In comparing the number of psychologists graduated from various institutions, we must take into account that one college or university has taught psychology five years, another, fifteen years; that one has only undergraduate courses in the subject, another, graduate courses, as well; that one has required courses, another, only electives; that one has, in proportion to its size, a large psychological faculty, another, a small faculty. These are factors that we ought not, perhaps, to expect to find in a broad interpretation of the statistical results, but they are worth noting, if only to show the limitations to which the figures are subject.

I have not meant, in commenting on Professor Cattell's statistics, to throw doubt upon the validity of the method. I should venture, rather, to suggest that the method be further elaborated and refined. The first report upon the returns leads me to think that elaboration and refinement will make the method still more useful. Subsequent reports may, however, prove that the modifications suggested were already under consideration.

NOCTURNAL EMISSIONS.

Knowing how much harm is often done to young men by the injudicious, nay criminal, advice of young physicians and others, not to speak of the iniquitous advertising quack, and on the other hand how much good is done by a frank assurance that the "symptoms" so often feared are purely physiological and rendered abnormal only by morbid introspection, I, with some natural hesitation, submit this article in the hope that it may be widely read and may do something toward mitigating a real social evil.

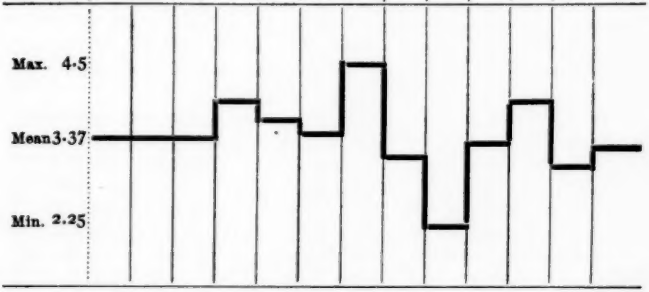
It is well known to physicians that a great deal of worry and misery is caused in young people by a lack of knowledge on their part, and on the part of their parents, regarding sexual matters. One great source of anxiety is nocturnal seminal emissions and the sexual or erotic dreams that commonly accompany them. These are regarded by ignorant youth and parents alike as a sign of depravity, and the consequences of this view are deplorable, physically, mentally and morally. Most physicians, and many persons who, like myself, are not physicians could give illustrations of the truth of this statement. Knowing this, and finding so much diversity of opinion regarding the subject, the present writer, a bachelor who has no record as a rake, undertook to record his own emissions—involuntary of course—and began to keep a record of them and of such accompanying dreams as he could recall early in October, 1895, when just turning thirty. This I have continued to do to the present (Sept. 30, 1903), namely eight years. As I know of no previous record of this kind and scope,[1] I think it worth publishing for the information of the physiologist, the psychologist and the physician, and indeed for all men, for it should have an interest for all whether married or unmarried.

I shall first give a table showing the number of emissions per month throughout the whole period, and then a curve showing the rise and fall of the number of emissions per calendar month on an average.

From the table it is evident that the number of emissions per year is pretty constant, varying between thirty-seven and fifty, or an average of between three and four per month

[1] While this article was printing, my attention was called to the following previous articles on the subject. Perry Coste: *Sexual Periodicity in Men*, in Havelock Ellis's Psychology of Sex, Vol. II, pp. 251-266; which gives average per year as 37. Nelson, Julius: *A Study of Dreams*, *Am. Jour. of Psy.*, May, 1888, Vol. I, pp. 367-401 (especially pp. 390-401); which gives curve of nocturnal emissions ("ekbole").

Year.	Jan.	Feb.	Mar	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Totals
1895										2	2	3	7
1896	4	4	2	6	5	5	5	7	2	2	6	2	50
1897	4	5	4	4	4	1	9	3	1	3	7	4	49
1898	3	4	3	2	3	3	5	2	3	3	6	2	39
1899	3	2	4	2	2	6	7	3	1	1	4	2	37
1900	3	2	5	2	4	5	1	2	3	7	1	4	39
1901	5	6	6	6	2	1	7	3	3	6	2	4	49
1902	3	3	3	5	5	4	2	2	3	3	2	3	38
1903	2	1	2	4	4	3	0	3	2				21
Totals,	27	27	27	31	29	28	36	25	18	27	30	24	329
Average,	3·37	3·37	3·37	3·88	3·62	3·5	4·50	3·12	2·25	3·37	3·75	3	41·12



throughout the time, except during the last year, when for the nine months recorded the average is a little less. This falling off may be due to the unusual demands made upon the system to repair an injury I received in February, which crippled me for several months. The exact average per month for the eight years is 3.43. That is to say the writer's normal or physiological sexual capacity is between three and four emissions per month. This capacity will no doubt vary greatly in different individuals. In some it will be greater, in others less than this, but there is decidedly a physiological limit. There is here a warning both to the unmarried masturbator and the married incontinent, who are equally foolish and blameworthy.

The average curve is also interesting as showing a spring and summer rise in the number of emissions and a corresponding fall, naturally following upon the seasonal welling up of life. Such a seasonal phenomenon is another link that unites us with the rest of organic creation.

The emissions, besides varying greatly in amount, are very irregular in time. The least interval between them was one day—not counting two taking place the same night, which

occurred, but very rarely. The interval of one day occurred in about 8.5 per centum of the cases. Forty-one per cent. took place in intervals of one, two, three or four days, fifty-nine per cent. with intervals of a week or less. Thirty per cent. took place in intervals of from eight to seventeen days, and but eleven per cent. with longer intervals. The longest interval between emissions was forty-two days. While both the amount and the frequency seemed to be influenced slightly by conditions of temperature, food, sexual excitement, etc., the influence was only temporary and made no difference in large periods of time. It has been said that urinating will relieve the pressure and quiet erection. I gave this a fair trial for a long period and found no difference in the net result, though sometimes an emission might be thereby postponed a day or so. Getting out of bed and walking up and down to allay erection was of similarly temporary effect. Illness, however, or poor condition, as also over-work and loss of sleep, certainly tended to make the emissions less frequent.

The time at which the emissions took place was generally in the early morning. Frequently I could not tell just when they occurred, as I knew nothing of it till I awoke. Generally, however, the emissions were accompanied by sexual dreams, which I was able to recall more or less clearly and out of which I sometimes awoke at the orgasm. Inhibition was very seldom possible and never more than temporary. I need scarcely say that the emission was never induced. I examined it once under a microscope and saw numbers of spermatozoa, proving it to be true semen.

The emissions, as I have said, were generally accompanied—or rather preceded—by dreams of a more or less erotic nature. Needless to say that legs and breasts often figured prominently in these dreams, the other sexual parts, however, very seldom and then they turned out to be male organs in most cases. There were but two instances of copulation dreamt. Girls and young women were the usual *dramatis personæ*, and, curiously enough, often the aggressors. Sometime the face or faces were well known, sometimes only once seen, somtimes entirely unknown. The orgasm occurs at the most erotic part of the dream, the physical and psychical running parallel. This most erotic or suggestive part of the dream was very often quite an innocent looking incident enough. As for example: While passing a strange young woman overtaken on the street, she calls after me some question. At first I pay no heed, but when she calls again, I hesitate whether to turn back and answer or not—emission. Again, walking beside a young woman, she said, "Shall I take your arm?" I offered it and she took it entwining her arm around it and raising it high—

emission. I could feel stronger erection as she asked the question. Sometimes a word was enough, sometimes a gesture. Once emission took place on my noticing the young woman's diminished finger-nails. Another example of fetichism was my being curiously attracted in a dream by the pretty embroidered figure on a little girl's dress. As an illustration of the strange metamorphoses that occur in dreams, I one night in my dream (I had been observing partridges in the summer) fell in love with a partridge, which changed under my caresses to a beautiful girl, who yet retained an indiscribable wild-bird innocence, grace and charm—a sort of Undina! Imaginative writers doubtless owe much to their dreams, in which the imagination has such free play. Sometimes the effort was made to inhibit the tendency to emission, but this was very seldom successful. Occasionally inhibition was due to the entrance (in dream of course) of a third person upon the scene. I did not find that the position of the body in sleep mattered much. Generally I was on my back, but sometimes on my side and rarely on my stomach. If I went to bed late very tired or unable to sleep much for thinking, the usual morning erection would generally not take place. Sometimes the erotic dream and emission left a strained or tired and unpleasant feeling behind; at other times I felt refreshed, relieved and the better for it. The erotic dream and emission frequently seemed to be the climax to a gradually increasing feeling of well-being or euphoria lasting some days.

Throughout these years my health has continued very good with rare exceptions, generally in spring, when I have suffered for short periods from tonsilitis; and my weight has varied but little.

LITERATURE.

AN EXAMINATION OF HÖFFDING'S THEORY OF RELIGION.

In his recent work on the Philosophy of Religion, Professor Höffding maintains the thesis that the essence of religion is a belief in the persistence of value. The indication of this thesis is not by any means the only purpose of the work, although it is a conspicuous feature of it. The present paper does not aim to review the work as a whole, but simply to examine the meaning and validity of this theory of the nature of religion.

That faith in the conservation of value constitutes the essence and ideal of all religion, is a proposition at once fascinating and baffling. It seems to bring religion at a stroke into line with physics and its doctrine of the conservation of energy. At the same time this very parallel makes the expression a difficult one to understand. Energy, from the point of view of the physicist, whatever the metaphysical may think of it, is an entity. It is clear enough what is meant, when the physicist tells us that the total quantity of energy remains constant under all changes of form. But we do not conceive of value as an entity. It is not a thing but a quality, and not even an objective quality at that, but the quality of certain things in their relation to beings possessed of an affective consciousness. In short, value is a ratio. What sense, then, can we attach to the phrase: conservation of value? We can best get at Professor Höffding's meaning if we look at some of the passages in which he uses it. The first preliminary statement of the theory is given in the introduction as follows: "If religious ideas have no value as knowledge, then the value which possibly they do possess, must consist in the fact that they bring to expression other sides of the soul's life than the intellectual. We must seek, then, a description of the religious life of the soul; particularly of the relation of the ideas to religious experience and religious feeling. In this way it will be shown that religion in its innermost nature has to do, not with the *understanding* of existence, but rather with the *valuation* of it, and that religious ideas express the relation in which actual existence, as we know it, stands to that which gives to life its highest worth for us. The kernel of religion, according to the hypothesis, to which we are impelled, is a belief in the persistence of value in the world (*ein Glaube an die Erhaltung des Wertes*). This belief appears in great representations of perceptual imagery in the popular religions, particularly in the highest of them. Likewise, in those men who stand outside every popular religion, such a belief as this can live even if it is not here set forth in definite form.

"The transition from the psychological side of religion to the ethical is brought about naturally by the question, what ethical value it has that the persistence of value is believed in. Ethics has to do with the production of values, and we cannot deny the possibility that the work of producing values leaves neither time nor strength to spend in thinking of their persistence." In another passage he says: "The religious problem is concerned not merely with the persistence of value in the human world. Not only a psychological, but also a cosmological question is presented—the question as to the relation to existence as a

whole of what appears to man of the highest worth. Is there a connection of values with the laws and forces of existence? Are, perhaps, even these laws and forces determined ultimately by the highest values? Or, perhaps, can we attribute to the concept of value no validity at all outside the sphere of human life?"

"The hypothesis which we shall later seek to establish," says Höffding, "rests upon this: that the principle of the persistence of value is the peculiar religious axiom which is expressed in various ways in the various religions and by the various religious standpoints." Again, he says, "If I undertake to show that the principle mentioned is the fundamental thought of all religion, this attempt will be none the less successful because it is manifest that no religion expresses the principle with clearness and consistency. It will suffice if an expressed need or expressed tendency to lay hold of this fundamental thought can be shown, so that the standard for estimating the worth of religions in their mutual relations, will be the degree in which they can express and carry out the principle. When the persistence of value is spoken of, it might be understood as if value did not disappear, but existence always contained value, whether it increases or decreases, or does both by turns. I take the expression persistence of value, however, in analogy with the expression persistence of energy, so that the principle asserts the continued persistence of value throughout all changes of form. It will therefore be necessary, perhaps, to distinguish between potential and actual value as we distinguish between potential and actual energy, and the one distinction is just as clear and appropriate as the other.

"The nature of a being determines its needs, and the needs determine what has value for it. Accordingly the religious axiom indicates the necessity that the character of a religion be determined by the nature and needs of its adherents. For one cannot believe in earnest in the persistence of value which he does not know in some measure from his own experience."

One of the first questions that naturally arise on considering this theory is: Does, then, a belief in the persistence of any sort of value constitute religion in Höffding's view? Is it some special kind of value, or is it value in general; value in the abstract that he means? That no determination of the nature of the value in question is given in the oft repeated formula, "Belief in the persistence of value," is indeed, confusing. In the course of the exposition, however, we find religious values differentiated from the other values in two ways. In the first place it is a belief in the persistence of the *highest* values which makes religion. It is not, however, the highest in any ideal or absolute sense, but simply the highest for us, and this is, perhaps, why Höffding does not put in the term highest in his formal statement of the principle. The Greek and the Greenlander may each have religion, according as he believes in the persistence of the value which is to him the highest, although the thing of highest worth may be eternal beauty for the one, and unlimited blubber for the other. Religion presupposes that man has some experience of value. The content of religion is always dependent upon man's experience, in particular the experience of what he has found worthful. "The values," Höffding says, "in whose persistence man believes, will be those which are to him the highest." And so the various religions may present the greatest differences, and yet each in its own way express a faith in the persistence of value. The opposition between the lower and the higher religion is of the greatest weight practically, but theoretically they are all the same. Lower and higher alike express in some measure the characteristic religious axiom—belief in the persistence of value.

Again, religious values have a certain derivative or secondary character as compared with other values, the primary values. Among our direct or primary values Höffding distinguishes two groups. One group of these values goes along with self assertion from its most elementary up to its most ideal forms. The other group of values is connected with devotion to beings, relations, and tasks, which reach out beyond the conditions of isolated self-assertion. To this second group belong the ethical, æsthetic, and intellectual feelings. Religion depends upon the existence of a third group of values, indirect or secondary values, which depend upon the values of the first two groups. The values of the third group spring from our interest in winning and preserving in existence the primary values. Existence appears to men as a battlefield on which the fate of values is decided. Existence is a great drama in which man is at once both a player and a spectator. Were he only a player, his entire energy and interest would be absorbed in his own part in the play. He would have no time or strength for, or interest in, the course of the drama as a whole. Were he, on the other hand, only spectator, his attitude would be purely intellectual or æsthetic. But if he is both, he will have values himself which are in conflict with the play. Besides his participation in it, the picture, also, which he forms of the course of the drama, will influence him and determine his attitude. He will feel himself drawn into the great order and stream of things with his entire inmost being, and for the sake of the highest values known to him, so that there will arise in him a most lively feeling of pleasure, or pain. In its immediate form this feeling discharges itself in utterances of hope and fear, of admiration and honor, of joy and sorrow. Such utterances are in all fields the simplest judgments of value that there are. "This feeling determined by the fate of values in the struggle for existence," says Höffding, "is the religious feeling." This feeling, accordingly, is determined by the relation of value to reality. This relation, as it appears to man, determines the value which man ascribes to existence. Accordingly, the religious judgments are secondary judgments of value. In comparison with the primary judgments of value, by which the first two groups of values are expressed, they are derived. So much for the differentation of the religious values from the other values. Another question of paramount interest in considering any proposed theory as to the nature of religion is: How does the theory in view square with the actual form which religion has taken in history? Höffding submits to a careful and somewhat extended analysis the typical manifestations of the religious consciousness. He finds in all his religious axiom present implicitly if not explicitly. Indeed, the axiom is rather an unconscious ideal which each embodies in its own way and in greater or less measure. As we breathe without knowing the physiological laws of respiration, as the common consciousness acts in accordance with the laws of the association of ideas without knowing them, so the religious consciousness unconsciously manifests the principle of faith in the persistence of value. While all manifest, or at least tend to manifest this principle in some degree, no actual religion expresses it in its ideal form. "For the complete carrying out and verification of the belief in the persistence of value," says Höffding, "it would be required that nothing should occur in the world process as mere means or mere possibility, to say nothing of a sheer hinderance, that, on the contrary, whatever possesses immediate value should have always at the same time immediate value, and that all hinderances should, at the same time, be means also." Höffding frankly states that no actual religion realizes the principle in this ideal form; one wonders if he is acquainted with Christian Science. One of the two fundamental syllogisms on which Mrs. Eddy builds her

system is, "God is all—God is good, therefore, all is good and there is no evil." This would seem to afford an almost perfect realization of Höffding's principle in the ideal form. Höffding shows a most extensive knowledge of the historical manifestations of religion, and of the biographies of the great religious personages. One wonders at times, however, how his views might be affected if he were to become equally familiar with the varied and abundant florescence of the religious consciousness on the free soil of America. In spite of the optimism of the principle of the persistence of value in its ideal form, Höffding admits that it does not escape the difficulty of the problem of evil. For even if it be granted that all hinderances and all opposition were necessary means to the development and preservation of the worthful, still the question would arise why any means at all should be requisite, why the worthful should not be and prevail immediately, why there should be any difference or strife between value and reality?

It might appear as if the hypothesis of the persistence of value would be quite inconsistent with any pessimistic form of religion, but such, Höffding argues, is not the case. The pessimist must admit that there is something of worth in the world, but that it is to be won only by hard fighting and unavoidable suffering. He fixes his attention, before everything else, on the struggle and suffering, but still he necessarily has a belief in the possibility that the worthful can be secured by persistent fighting and suffering. No religion and no philosophy maintains an absolute pessimism. Buddha points to the possibility tf attaining nirvana which is not at all to be identified with annihilaoion. Schopenhauer sees escape from disharmony in artistic and scientific activity, in sympathy and religious asceticism. An absolute pessimism would be valid only for those beings which some religions damn to eternal pain. But this eternal pain is not the entire content in any religion. The effort is always made to show that the highest values are to be attained not only in spite of, but even because of the eternal pain in some part of existence. The complete opposite of all religion would be neither optimism nor pessimism, but, on the contrary, neutralism; the view of things, according to which all valuation beyond the field of human conduct falls away, and man regards the course of the great world process only as a spectator, feeling himself convinced of absolute indifference towards all that men call worth. But such indifference is, perhaps, more theoretical than possible. Even the mere spectator would feel some intellectual, even if no ethical interest. It would endow existence with some value that it gives joy to his own understanding or contemplation, and his view of the world would be colored involuntarily by the progress or retrogression of such intellectual and æsthetic values.

In this brief survey of Höffding's Theory of Religion, I have endeavored, by means of quotation and paraphrase, to let the theory speak for itself. Time permits only a brief indication of the criticisms that might be offered. In the first place, the analogy suggested between conservation of value and conservation of energy, seems somewhat too far fetched. It may be possible, however, to attach some meaning to this analogy, even though value be a ratio, and energy an entity. A ratio may be conceived as constant even though the objects related are constantly undergoing change. The analogy is helped out by noting the existence of potential values [the pause in music is instanced as an example of a potential value], and the analogy is carried further by noting the transformation which values undergo—as, for example, it frequently occurs that a value which is first mediate comes in time to be immediate. But when we come to the quantitative aspect which plays so important a part in the doctrine of the conservatism of energy, the analogy seems to break down altogether.

The notion of a quantum of value in the universe, and the permanent conservatism of that quantum is wholly elusive. The notion takes form at all only by hypostalising value. Again, we may question whether the principle of the conservation of value offers anything distinctive of religion. We may admit Höffding's success in showing that all religions involve this element, but does not all intellectual effort in whatever sphere, involve a faith in the continuance of values, and involve some sort of conception of the relation between value and reality in the realm of experience, as well as in the transcendental realm?

A belief in the maintenance of empirical values cannot constitute religion. It is only when our faith carries us into the transcendental realm, or, if you please, when we consider the relation of value or of ourselves to reality as a whole, that we come into the religious attitude. Again, granting that a faith in the persistence of value is a universal aspect of religion, we may still ask whether, after all, it can properly be regarded as the *essence* of religion. It seems to me there is a confusion here of effect with cause, or, so to speak, of attribute with substance. Tiele has pointed out, on the basis of his studies in the science and history of religion, that the characteristic of religion subjectively is adoration. Recall, now, Matthew Arnold's formula for the objective side of religion, "The power, not ourselves, that makes for righteousness." Combining these two we may say religion is our adoration of the "Power, not ourselves, that makes for righteousness." Undoubtedly, one outcome of such adoration is faith in the persistence of moral worth, but this confidence is rather a product of religion than religion itself. Again, Höffding's hypothesis fails to account for the moral dynamic in religion. Such a view might answer for certain quiescent, mystical, contemplative forms of religion. But it fails utterly to account for the moral force so characteristic of the active types of religion.

Such a view, further, fails to give religion its place with respect to the other aspects of our spiritual life. Just as, theoretically, God is the principle of unity in the world, so, practically, religion should be the principle of unity in life. We must appeal to our transcendent interest to preserve the proper balance in our several empirical interests. Or, as we may state it, if religion is to be of any value, it must set a criterion for values. In answer to a similar criticism which I made in the review of the work published last March, in the Philosophical Review, Professor Höffding, in a private letter, urges his view that it is ethics whose business it is to furnish a standard of values. But while ethics may properly afford the standard for appraising the value of conduct, we must call upon religion for the standard by which to appraise the relative worth of our several types of values themselves, such as the sensuous, the intellectual, the æsthetic, and the moral. Which of the various kinds of value possible to us shall we seek to realize, and in what proportion? To this practical question, where religion, if anywhere, ought to give us help, a belief in the conservation of value seems to be of no service. If all influence in practical life is to be denied to religion (and this is, perhaps, just what Höffding means to say), then religion is reduced to a merely subjective attitude of the æsthstic type toward reality as a whole. This view is borne out by the view of the future development of religion which Höffding founds upon his theory of its essence. Religion, he thinks, is to become more and more a form of poetic symbolism.

In conclusion, I will call attention to what seems to me one manifest excellence in Höffding's theory. There are two directions in which one may look to find the essence of any group of phenomena that constitute a progressive series. One way is to eliminate all the differ-

ences and seek the common element that is left. This is Herbert Spencer's method by which he finds the essence of religion to consist in a recognition of the inscrutable mystery of things. The second way seeks the essence in the idea that is successively, progressively, but perhaps never completely manifested in the series. The first method attempts to explain the higher by the lower—the second method finds the explanation of the lower only in the higher. The first seeks an elemental essence; the second, an ideal essence. The first might be denominated the logical method; the second, the biological method. It is a pre-eminent merit in Höffding's work that he has adopted the second method. Whatever may be the final estimate of his hypothesis, he is worthy of great credit for his effort to determine the ideal essence of the religious consciousness.

F. C. FRENCH,
Lincoln, Nebraska. University of Nebraska.

Heredity and Social Progress, by SIMON N. PATTEN. New York, The Macmillan Co., 1903. pp. vii, 214. $1.25.

Modern thought accepts the seeming paradox that evolutionary progress owes its origin to adverse conditions of environment. Thus, current biology and economics hold that the moulding factors of evolution are operative chiefly under conditions which are unfavorable to the life of the organism. And current religion teaches that morality thrives in adversity and decays in prosperity. Professor Patten maintains that this conception is fundamentally false. The main argument of his book supports the counter thesis: Progress starts from a surplus, not from a deficit. The acquired characters of one environment create a surplus of energy which evokes new characters; these in turn impel their possessors to seek another environment, in which the maximal benefit of the improved equipment may be reaped.

But how are the new characters evoked? And how do they become permanent possessions of the race? The accumulated surplus is made up of perishable goods only. To make progress secure, this temporary surplus must be transformed into permanent conditions or into mental traits. If such a transformation is impossible, progress can never be more than merely temporary. If it is possible, not only will a permanent progress be assured, but the process of the transformation must portray the course of social progress.

Professor Patten finds that the transformation actually does take place, and in the following manner. The successful artisan provides more food, leisure and protection for his children. Brought up under these conditions, the children acquire greater vitality, increased energy and a fuller development of their natural qualities. The new characters thus acquired beget a disinclination to the humdrum occupation of the father. New environments and new occupations are sought, in which the acquired characters may be utilized. The children become physicians, lawyers and clergymen; and move into a higher class in the social scale.

The essential point, then, in all social progress is the creation of a social surplus. This surplus is not permanent in form but is constantly in circuit, both biologically and economically, always disappearing but ever reappearing in new form. Its successive phases recur in the following order: First there is produced a surplus of energy through the employment of acquired characters; this surplus is next expressed in secondary characters; then a use is discovered for these secondary characters, in which the whole species can share; and finally the species moves into a new environment where the secondary characters are necessary.

The characters acquired by an individual or a generation, are not transmitted directly to its descendants. They gradually become in-

grained in the social fabric and are perpetuated in customs, habits and traditions. But the characters themselves must be reacquired by each succeeding generation. Then, too, it is found that the gradual acquisition of new characters is attended by a one-sided development of the individual. It becomes the function of education to restore uniformity; this it attempts to do, not by eliminating the weak, but by developing and strengthening them.

In the lower organisms, on the contrary, progress is made by means of elimination. But in human society differentiations are conserved and harmonized by a division of labor or of function. Nor are human contests fought out to the death; they are usually decided within the domain of wealth. One party yields when its property is endangered or destroyed, as in modern warfare, or when its material resources are shown to be inferior, as in industrial struggles between nations or classes. Hence, in modern times at least, natural selection acts upon man only indirectly, through his material equipment.

Professor Patten makes a good deal of emotion as a factor in progress. When a herd of deer is attacked by lions the emotion aroused by the slaughter impels the survivors to change their habitat, and possibly their habits and food as well. Emotion narrows choice, limits the range of food, and consequently causes an arrest of development. In the case cited, the adjustment to new conditions, ushered in by the emotive factor, occurs much more rapidly than could be effected by natural selection. National disasters act as a similar check to human progress, in virtue of their emotional effect.

The author assumes, as a matter of course, a thoroughgoing parallelism between biological and economic evolution. The result is that the discussion not infrequently degenerates to mere argument from analogy. Many of his biological conceptions are weird in the extreme: *e. g.*, his conception of the brain as an inclosed ovary, male in woman and female in man (pp. 105 ff.), and his theory of "reduction." Nor is his psychological doctrine less startling; *e. g.*, his treatment of memory, emotion and visualization (pp. 95 ff.). But these lapses do not invalidate his general position. Indeed, these parts of the discussion might have been omitted from the volume without serious loss.

Though the argument fails at times to carry conviction, the treatment has at least the virtue of originality, and will, doubtless, stimulate further discussion. J. W. BAIRD.

The Virgin Birth of Christ. An historical and critical essay by PAUL LOBSTEIN. Translated into English by Victor Leuliette. Edited by W. D. Morrison. G. P. Putnam's Sons, New York, 1903. pp. 138.

This remarkable and interesting little work is to be welcomed in its plain English dress. The author's general conclusion may be stated in his own language if the Virgin Birth of Christ "ceases to remain a real fact, it stands out as the characteristic creation of the faith of the church." "In the light of this two-fold investigation the tradition of the miracular birth of Jesus vanishes away, or rather resolves itself into a myth created by popular devotion and destined to explain the Divine Sonship of Christ by his supernatural generation. Thus viewed, the conception of our two evangelists is an important landmark in the development of Scriptural Christology. Every attempt made to reconstruct on a dogmatic basis a notion now forever overthrown by exegesis and criticism, has but betrayed the inextricable contradictions of popular orthodoxy. We are, therefore, doing a valuable service to faith by confining the traditional doctrine to its religious import. By laying bare the imperishable truth contained in the Christian symbol, we are remaining faithful to the mind of Jesus,

who bases his own unique communion with God not upon the physical miracle of his supernatural birth, but upon the sovereign election and perfect revelation of the Heavenly Father. If we set aside the dogma which the school has deduced from the stories of Matthew and Luke, it is that we may subscribe with absolute sincerity and joyful assurance to the Johannine saying, echo of our experience and confirmation of our faith: 'Jesus said unto them: I am from above; I am not of this world.'"

The Unwritten Sayings of Christ, by C. G. GRIFFINHOOFE. Edward Arnold, London, 1903. pp. 128.

This is an interesting and somewhat popular account of the words imputed to Jesus, not only in the Apocryphal Gospels but in the writings of some of the early fathers, including the newly discovered logia. It is evident that the writer believes that in many of these we have substanially the very expressions of our Lord.

Les limites du connaissable. La vie et les phénomènes naturels, par FÉLIX LE DANTEC. F. Alcan, Paris, 1903. pp. 237.

This volume is made up of various recent essays of the author, the first and most important being a characterization of the place of life in natural phenomena. Repose is an illusion in the world of brute matter. Chemical reaction, action at a distance, and vibrations at least are continuous. The dimensions of life and its lower limits of size are next discussed, and the forms of movement that the human mind can know. Other briefer chapters deal with Grasset's conception of the limits of biology, the divine, retrograde movement, evolution and its apologists, and knowledge of the future.

Animal Education. An experimental study of the psychical development of the white rat, correlated with the growth of its nervous system, by JOHN B. WATSON. University Press, Chicago, 1903. pp. 122. Price, $1.25.

This is, perhaps, the best study of many that have lately been made of the white rat. The experiments, themselves, are careful, but the significant feature of the work is a comparison of the psychic development of the white rat at different ages, and the correlation of its stages of development with the growth of the inner structure of the brain. Although no momentous conclusions are drawn, close parallelism of psychic growth with that of medullation fails to appear.

Willensfreiheit und moderner psychologischer Determinismus, von ANTON SEITZ. J. P. Bachem, Köln a. Rh. pp. 62.

Die Energie und Entropie der Naturkräfte mit Hinweis auf den in dem Entropiegesetze liegenden Schöpferbeweis, von R. SCHWEITZER. J. P. Bachem, Köln a. Rh. pp. 59.

An Introductory Study of Ethics, by WARNER FITE. Longmans, Green and Co., London, 1903. pp. 383.

This work treats first of scope, method and problems; then of the various forms of hedonism, empirical, scientific, theoretical; their relations to common sense; then of intuitionism or the ethics of conscience; rationalism or the ethics of principle; self realization or the ethics of purpose; idealistic, social theory; idealism as a system of philosophy; idealism and common sense. The third part discusses the moral situation, or the metaphysical from an evolutionary standpoint, and the practical, moral attitude, with illustrations. The book has few references to literature.

Neurologic Progress and Prospects, by F. W. LANGDON. Chicago, 1903. pp. 19.

This author gives a very interesting account of the scientific, and especially the clinical and practical progress of recent years in neurology.

The Wellcome Physiological Research Laboratories. Founded 1894. Walter Dowson, Director. London.

This laboratory is admirably equipped, and this little booklet gives us a very interesting picture of it, with a good account of the work which is thought to be desirable in such institutions.

Grundzüge der physiologischen Psychologie, von WILHELM WUNDT. Fünfte völlig umgearbeitete Auflage. Dritter Band. W. Engelmann, Leipzig, 1903. pp. 796. Unbound, M. 14. Bound, M. 17.

This volume treats of time concepts, feeling concepts and passions, will processes, consciousness and the lapse of concepts, psychic connections, anomalies of consciousness, scientific preconceptions of psychology and its principles. An index is to follow.

Das Problem der Gegebenheit, von PAUL STERN. Grundprobleme der Philosophie. Bruno Cassirer, Berlin, 1903. pp. 79.

In the first part the writer gives us a critique of false presentations and their consequences in empiricism, sensualism, etc. The second part describes the problem of building up the world and mechanical psychology, with special reference to evolution and psychological metaphysics.

The Free-Will Problem in Modern Thought, by WILLIAM HALLOCK JOHNSON. The Macmillan Co., New York, August, 1903. $.75. pp. 94.

Free will is discussed, first with relation to the psychophysic question, then as related to evolution. The third section is on the consciousness of freedom; in the fourth part it is considered as an ethical postulate; and, lastly, its relations to theology are discussed.

The Edwardean. Devoted to the History of Thought in America. Ed. by Wm. H. Squeres, Hamilton College, New York, Oct., 1903. Vol. I, No. 1. $2.00 per yr.

This new quarterly marks the two hundredth anniversary of the birth of Jonathan Edwards, and is to be devoted to the history of thought in America. The articles are glimpses into Edwards's life, the seventy resolutions, Edwards as a thinker, some estimates of his work, and his metaphysical foundations.

St Anselm. Pros. Logium: Monologym. An Appendix in Behalf of the Fool, by GAUNILON and CUR DEUS HOMO. Tr. by Sidney Norton Dean. Open Court Publishing Co., Chicago, 1903. pp. 288. $.50.

Queries in Ethnology, by ALBERT GALLOWAY KELLER. Longmans, Green & Co., New York, 1903. pp. 77.

This little vest pocket manual consists of eight hundred and sixty-nine points or suggestions for those observing primitive people. They are classified under maintenance, perpetuation, gratification, religious and superstitious ideas and uses, the societal system, and contact and modification.

Dissertations on Leading Philosophical Topics, by ALEXANDER BAIN. (Mainly reprints from *Mind*). Longmans, Green & Co., New York, 1903. pp. 277.

These papers, mainly reprints from *Mind*, cover the more important

of the author's incidental contributions to philosophy during the last twenty-five years. Some of the most interesting of these are on the scope of anthropology and its relations to the science of mind, the mutual help of introspection and psychophysic experiment in psychology, physiological expression in psychology, is there such a thing as pure malevolence, definition and problem of consciousness, etc.

Lao-tze's Tao Tech King. The Canon of Reason and Virtue, tr. from the Chinese by DR. PAUL CARUS. Open Court Publishing Co., Chicago, 1903. pp. 138. $.50.

The Open Court Publishing Company is rendering a real service to students of philosophy, religion, and psychology, by placing such standards as these within the reach of the public at such moderate cost. The last is number 55 in the series.

Zur Psychologie des ästhetischen Genusses, von DR. G. WERNICK. Verlag von Wilhelm Engelmann, Leipzig, 1903. pp. 148.

The author first makes some general observations concerning æsthetic effect, then studies the laws of association, reproduction and sensation in their relations to æsthetic feeling, and these three points constitute the substance of his work.

The Philosophy of Auguste Comte, by L. LEVY-BRUHL. Introduction by Frederic Harrison. G. P. Putnam's Sons, London, 1903. pp. 363. $3.50.

We have here an admirable condensation of the four parts or books of Comte's positive philosophy. It was a bold undertaking to compete with the masterly condensation of Harriet Martineau, but it was certainly desirable to have a briefer epitome, and this we have here. It is enough to say that it does not suffer by comparison with Harriet Martineau's work.

The Nature of Man. Studies in Optimistic Philosophy, by ÉLIE METCHINKOFF. English translation edited by P. Chalmers Mitchell. G. P. Putnam's Sons, New York, 1903. pp. 309. $2.00.

The writer first treats of disharmonies in man's nature, then of the attempts to diminish the ills arising therefrom, and in the last part résumés what science is able to do to alleviate the disharmonies of the human constitution, with a characterization of Pasteur and his work, in whose institute the author's own sphere of work lies.

Studies in Logical Theory, by JOHN DEWEY and others. The Decennial Publications. 2nd ser. Vol. XI. Univ. of Chicago Press, Chicago, 1903. pp. 388. $2.50.

This volume contains four papers by the editor, and seven by other people, mostly professors elsewhere, on various logical topics. The book not only has high intrinsic value, and should be indispensable to every logician, but it is an interesting presentation of the work accomplished by one of our leading academic professors and those who have formerly been his students.

Le Movement, par R. S. WOODWORTH. Tr. par Mlle. le Dr. Elvire Samfiresco. Octave Doin, Paris, 1903. pp. 421.

In the first part the author discusses the perception of bodily movements, and in the second part their production. On the whole it is a valuable summary of the condition of our experimental and theoretical knowledge on the subject, with an excellent bibliography, index, etc.

Hypnotism. Its history, practice, and theory, by J. MILNE BRAMWELL. Grant Richards, London, 1903. pp. 477.

Of these fifteen chapters the best are those devoted to the experi-

mental phenomena of hypnosis and its uses in medicine, and best of all a thorough discussion of hypnotic theories. The author's own views here are sound and scientific, and he takes a wholesome stand against the mysticism which has lately crept into the conceptions and discussions of this topic.

L'Annde Psychologique, par ALFRED BINET. Vol. 8. Schleicher Frères, Paris, 1902. pp. 757.

The first 389 pages are devoted to original articles, sixteen in number; then follow the digests and discussions ending with page 583; the rest of the volume being devoted to titles.

Transactions of the Congress of American Physicians and Surgeons. Sixth Triennial Session held at Washington, D. C., May 12th, 13th, and 14th, 1903. Published by the Congress, New Haven, 1903. pp. 243.

The Critical Review of Theological and Philosophical Literature, edited by S. D. F. Salmond. Vol. XIII, March, 1903. No. 2, pp. 192. Williams and Norgate, London.

Die Einwanderungspolitik und die Bevölkerungsfrage der Vereinigten Staaten von Amerika, von R. KUCZYNSKI. L. Simion, Berlin, 1903. pp. 35.

The Relations of Structural and Functional Psychology to Philosophy, by JAMES ROWLAND ANGELL. The University Press, Chicago, 1903. pp. 21.

Polydactylism in Man and the Domestic Animals, with Especial Reference to Digital Variations in Swine, by C. W. PRENTISS. (Bull. of the Mus. of Comp. Zool., Vol. XL.) Cambridge, Mass., 1903.

Dr. J. W. Wherry (Iowa Medical Journal) attempts to answer the question, "Why have there been so few results from original research in State hospitals for the insane?" This, he thinks, due largely to politics which makes careers so uncertain that even the position of assistant is felt by those who hold it to be tentative, and to the fact that specialists have almost ceased to study insanity *per se* and drifted away from the prime intention of hospitals, which was to shed light upon its origin, prevention and cure. He declares that insanity is not a disease but a condition expressing a mental attitude. Even could we prove that molecular brain change is its cause we should still need to seek a cause of the cause. Cords of histories of cases have been written and even printed, but scientific research begins where history ends. It is not even proven that insanity is a primary brain disease, because function, at least in many cases, seems to precede structural changes. Masses of literature have been written in the hope of some imposing future genius who would find something valuable in it. Bacteriology has little or nothing to do with insanity. Physicians have been led captive of late by the theory that insanity is the result of a brain disease. This has caused an infatuation and has minimized attention to psychic abnormalities, which have no physical symptons or correlates, and has led to the abandonment of the subject by psychological minds.

The Theory of Advertising, by WALTER DILL SCOTT, Ph. D. Small, Maynard & Co., Boston, 1903. pp. 240.

Mr. Scott, who is the director of the Psychological Laboratory at Northwestern University, has in this book made an application of

psychological principles to the theory of advertising. Nearly all of the chapters were originally published in Mokin's Magazine, and have now been collected and given to the public in book form. The author has analyzed a large number of advertisements and points out how their success or failure depends upon the degree in which they follow or deviate from psychological principles. Mr. Scott's theories have, in several cases, been practically tested by business firms, and have proved successful. The book is primarily designed as a guide in advertising, but is interesting to psychologists as a practical application of familiar principles.

Professor Strong on the Relation Between the Mind and the Body, by MORTON PRINCE. The Psychological Review, Vol. X, No. 6, Nov., 1903. pp. 658.

Mr. Prince claims to have anticipated Professor Strong as early as 1888, in his solution of the theory of the relation of the mind to the body. In this discussion he used the following words:

"In other words, a mental state and those physical changes which are known in the objective world as neural undulations, are one and the same thing, but the former is the actuality, the latter, a mode by which it is presented to the consciousness of a second person, *i. e.*, to the non-possessor of it.

"The real question is, not regarding the transformation of matter into mind, but how one state of consciousness comes to be perceived as another state of consciousness, or how a subjective fact comes to be perceived as an objective fact; how a feeling comes to be presented to us as a vibration."

"Physical changes (the thing-in-itself) occurring in a foreign body, as a piece of iron, though giving us our experience of it, must be absolutely unknown to us. Physical changes occurring in our brains are clearly known to us; they are our thoughts, our sensations, and our emotions."

"The common expression that 'every state of consciousness is accompanied with a molecular change in the substance of the brain,' . . . . must be regarded as unfounded and as leading to great confusion and misconception. A feeling is *not* accompanied by a molecular change in the same brain; it is 'the reality itself of that change.' . . . . You cannot correctly say that a feeling is accompanied by a molecular change in the same organism, because this implies two distinct existences, and leads to all the fallacies of materialism.

"The parallelism is between your consciousness and my consciousness of your consciousness, or, what is the same thing, between the consciousness in you and the picture in my mind of neural vibrations."

Dr. Prince regrets that Professor Strong had overlooked his own book, but welcomes him as a newcomer to the pan-psychic doctrine of which his work is a "capital restatement."

Evolution and Adaptation, by THOMAS HUNT MORGAN. The Macmillan Co., New York, 1903. pp. 470.

This volume is in a very different and wider field from that of the author's work on regeneration, and to many it will be a disappointment. Much in the early chapters is certainly a little tiresome and familiar. The statement of Mendel's law and of the mutation of De Vries, of Nägeli's perfecting principle, is very convenient for non-biologists who have not read the original. So is the statement of the doctrine of tropisms and instinct. But in the latter the author appears to know very little of this immense field, and to psychologists, at least, who, perhaps, first turn to this chapter, the disappointment will be most felt. It is convenient, however, to have so many of the prob-

lems of evolution stated in a concise way between two covers. Perhaps the most exasperating feature of the book is the almost complete absence of references.

L'Origine dei Fenomeni Psichici e il loro Significato Biologico, by G. SERGI. (Biblioteca di Scienze Moderne. No. 14.) F. Bocca, Torino, 1904. pp. 366.

Professor Sergi here gathers up in a very interesting way his own essentially evolutionary conceptions of the origin of the soul. Sensibility, movement, memory, are essentially biological functions, and the psychic heredity is as real as that of organism. He gives a valuable account of the psychic function of nutrition and instinct, and traces the development down to the social and religious consciousness, with plenty of space given to the abnormal side of the topic.

The Religious Sense in its Scientific Aspect, by GREVILLE MACDONALD. Hodder & Stoughton, London, 1903. pp. 243.

This work consists of three lectures, the Religion of Service, of Renunciation, of Freedom. The writer draws the illustrations from cuts of the foraminifera, the spongilla, the Venus flower-basket, the daisy, the guilder rose, pearly nautilus. While it cannot be said to add to our knowledge, or to be aggressive or original, it leaves a very pleasant impression from its easy and complacent style.

Gefühl und Bewusstseinslage, von JOHANNES ORTH. Reuther & Reichard, Berlin, 1903. pp. 131.

The writer urges that feeling is an independent psychic phenomenon and attempts to distinguish it from sensation and will, argues that pleasure and pain must still be regarded as feelings, denies specific will feeling, holds that excitement, tension, and rest are not true feelings, postulates a revision of the idea *affekt*, and urges that doubt is not a feeling.

The Philosophy of Hobbes in Extracts and Notes Collated from His Writings, by FREDERICK J. E. WOODBRIDGE. The H. W. Wilson Co., Minneapolis, 1903. pp. 391.

Every student of philosophy will heartily welcome this text of selections, and all those teachers who refer their classes to original authorities in the history of philosophy, will be glad to have the most readable parts in convenient form.

A Non Surgical Treatise on Diseases of the Prostate Gland and Adnexa, by GEORGE WHITFIELD OVERALL. Marsh & Grant Co., Chicago, 1903. pp. 207.

Man's Immortality and Destiny, by ROBERT P. DOWNES. "Great Thoughts" Office, London. pp. 180.

This is one of those books from which the publisher has very carefully scored away everything that can indicate its date. Whether it is a very old book now reprinted, or brand new, is impossible to tell. The theology of it reeks with age. It is poetic and sermonesque.

ERRATA.

Vol. XIV, p. 422 (Commemorative Number, p. 158) in Protocol for Fig. 2:

Question: for	34 mm.	*read*	44 mm.
" "	3502σ.	"	4532σ.
Error: "	-470σ.	"	-1500σ.

N. B.—Values given in Table I, Subject S. (p. 425 or 161), are unaffected by the mistake, the protocol as corrected being a part of the data for the *last* record there given.